Drinking and depression in Norwegian doctors: a 15-year longitudinal study

Kjersti Støen Grotmol

Department of Behavioural Sciences in Medicine

Faculty of Medicine

University of Oslo

Oslo, 2012

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 1411

ISBN 978-82-8264-400-6

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika publishing.
The thesis is produced by Unipub merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

ACKNOWLEDGEMENTS

This thesis was completed at the Department of Behavioural Sciences in Medicine, Faculty of Medicine, University of Oslo, and was funded by the Norwegian Medical Association and the Norwegian Research Council. The study on which the present thesis is based was initiated in 1993 by Professors Per Vaglum and Øivind Ekeberg. Work on the thesis was carried out from January 2007 to January 2012.

First, I would like to thank my principal mentor, Professor Reidar Tyssen, for his excellent supervision. His dedication to research on the health of medical students and physicians, combined with great enthusiasm and encouragement, has been very inspiring. Professor Tyssen has been available to address my enquiries at any time, and he has been very helpful, with creativity and determination, in all stages of my work. The clear minds of my two other supervisors, Professors Per Vaglum and Arnstein Finset, have also contributed substantially. I have also received invaluable help with methodological issues from Professor Torbjørn Moum. Professor Olaf Aasland at the Research Institute of the Norwegian Medical Association not only has played a central role in the design and implementation of the survey on which this thesis is based but also has pointed out the relevant literature and shared data on physicians in the course of my research.

I am also grateful to all of the members of our research group and my other considerate, clever colleagues in our department. My thoughts have been clarified more than once by the Tuesday research meeting, and the informal chats in the corridors have been very important to me.

I am also indebted to my family. In particular, my father Tom has been extremely helpful on the side-lines, sharing his extensive research experience with his at times frustrated daughter. My dear husband, Anders, has supported me throughout my work, expressing nothing but firm belief in my ability. Thank you to my two small children, Ingrid and Magnus, for providing a fun and challenging home environment, and for being a great source of inspiration, particularly when referee statements may have seemed overwhelmingly negative. My mother Randi, my sister Yngvild, and my many good friends have also been great conversation partners during these years.

TABLE OF CONTENTS

PREFACE

I am the daughter of two physicians, and there are many physicians in our extended family and circle of acquaintances. While growing up, I clearly saw the two-sidedness of what I believe is the experience of many physicians: their enthusiasm and love of their work, but also how they are affected by long work hours, lack of sleep, and the many dreadful situations they encounter. For a long time, I considered physicians to be strong, resourceful individuals, perhaps better equipped than others to tackle the many challenges of life. This impression has gradually been replaced by the realization that they, like the rest of us, are vulnerable.

My interest in alcohol problems grew with my experience as a psychologist at Lade Behandlingssenter, an institution for women with substance abuse problems in Trondheim. Because the majority of the women I met struggled with depressive symptoms as well as substance abuse, I became fascinated by these two conditions that so often go hand in hand. Why? Do these women drink to cope with their symptoms? Do they get depressed because they drink? What other factors have contributed to their problems?

When I read the proclamation/announcement of a PhD scholarship on this exact topic, *Alcohol problems and depression among physicians*, I felt it was perfect for me. It combined my professional and personal interests perfectly because I was sincerely interested in which of these intelligent, admired, and high-achieving physicians will, against all odds (or perhaps not?), end up struggling with these problems.

SUMMARY

Physicians are at greater risk of suicide than academics and other health professionals, and report lower life satisfaction than the general population. Alcohol problems and depression, individually and together, increase the risk of suicide. Several studies have identified high levels of depressive symptoms among interns and physicians in their early postgraduate years. Alcohol is the drug of choice among the physicians who seek treatment for addiction, and in combination with alcohol, "mood disorder" is the most common dual diagnosis for physician patients. Mental-health problems can also interfere with a physician's care of his/her patients.

Despite these troubling observations, long-term cohort studies of depression and alcohol problems among more-established physicians are lacking. To provide more detailed information about the prevalence, courses, and individual risk factors of these two conditions, a longitudinal follow-up study of two nationwide Norwegian cohorts of medical students was launched in 1993. All students who commenced medical school in Norway in 1993 and those who graduated in 1993 and 1994 have been followed up in five stages, with the last follow-up in 2008, when the oldest cohort of physicians was 42 years of age. Of all the long-term follow-up studies on physicians undertaken to date, the present Longitudinal Study of Norwegian Medical Students and Doctors (NORDOC) is the most representative and has the longest follow-up period.

The first paper included in this thesis supports the hypothesis that positive expectancy about the tension-reducing effects of alcohol, measured four years after graduation, can predict hazardous drinking six years later. This effect was not mediated by the actual use of alcohol to cope, measured at the same time as alcohol expectancy. The prevalence of hazardous drinking did not change between the two observation points and was reported by 9% of subjects four years after graduation and by 8% six years later. Both alcohol expectancy and hazardous drinking were more prominent among men, but hazardous drinking among women may have been under-estimated.

In the second paper, suboptimal perceived parenting by the mother, measured at the end of medical school, predicted severe depressive symptoms several years later (four and 10 years). This relationship, which was stronger in men, was partly explained by low self-esteem. Severe depressive symptoms did not change for the whole group between the two observation points, and there were no sex-based differences in prevalence.

In the third paper, additional possible risk factors for severe depressive symptoms were investigated: parental bonding, age, sex, personality, perceived medical school stress,

and perceived clinical skills and competence. Of these, young age and neuroticism were found to be risk factors. A decline in the prevalence of severe depressive symptoms was noted throughout the follow-up period. No sex-based differences were identified in the prevalence of or risk factors for severe depressive symptoms.

Together, these results show that individual risk factors for both depressive symptoms and hazardous drinking can be identified early in a medical career. Although there was a decline in severe depressive symptoms after the doctors left medical school, there was no such decline in hazardous drinking, the prevalence of which remained the same as that observed at medical school, even during the postgraduate years. Measures should be taken and interventional efforts applied to combat both depression and hazardous drinking among doctors early in their medical careers, even at the undergraduate stage. Further studies are required to determine the relative contributions of individual and contextual factors.

ABBREVIATIONS

NORDOC = The Longitudinal Study of Norwegian Medical Students and Doctors

GSE = General Self-Esteem

PBI = The Parental Bonding Instrument

GHQ-28 = The General Health Questionnaire

GEE = generalized estimating equation

AEQ = The Alcohol Expectancy Questionnaire

PMSS = Perceived Medical School Stress

BCI = The Basic Character Inventory

AUDIT = The Alcohol Use Disorder Identification Test

LIST OF PAPERS

Paper I

Kjersti s. Grotmol, Per Vaglum, Øivind Ekeberg, Tore Gude Olaf G. Aasland and Reidar Tyssen

Alcohol expectancy and hazardous drinking: a 6-year longitudinal and nationwide study of medical doctors

European Addiction Research 2010 16:17–22.

Paper II

Kjersti S. Grotmol, Øivind Ekeberg, Arnstein Finset, Tore Gude, Torbjørn Moum, Per Vaglum and Reidar Tyssen

Parental bonding and self-esteem as predictors of severe depressive symptoms: a 10-year follow-up study of Norwegian physicians

The Journal of Nervous and Mental Disease 2010 198:22–27.

Paper III

Kjersti S. Grotmol, Tore Gude, Torbjørn Moum, Per Vaglum and Reidar Tyssen

Risk factors at medical school for later severe depression: a 15-year longitudinal, nationwide study (NORDOC)*

Submitted for publication

*The Longitudinal Study of Norwegian Medical Students and Doctors

1 INTRODUCTION

This thesis examines the individual risk factors for two relatively common and important mental-health-related problems experienced by physicians: depressive symptoms and hazardous drinking. A mentally ill physician may represent a serious risk to his/her patients(1), yet both medical students and physicians are less likely than the general population to seek and receive appropriate treatment (2;3). Depression is prevalent at least among young physicians (4), and alcohol problems can lead to further health problems among doctors who already suffer from depression (5;6). The co-morbidity of alcohol abuse and depression amplifies the already elevated suicide risk among physicians.

To identify individual risk factors for depression and hazardous drinking, longitudinal studies of large representative samples are required. Such research is scarce because to date, these issues have primarily been studied with cross-sectional designs. The few published prospective studies are characterized by short follow-up periods and low response rates. Therefore, in 1993, two professors at the Department of Behavioural Sciences in Medicine at the Faculty of Medicine, University of Oslo, Per Vaglum and Øivind Ekeberg, undertook the largest prospective representative study of physicians' mental health to date, the Longitudinal Study of Norwegian Medical Students and Doctors (NORDOC). The present thesis is based on three studies conducted on the Young Doctor Cohort, one of two samples that together constituted the NORDOC study population. The Young Doctor Cohort comprised all Norwegian medical students (631 in total) who graduated in 1993 and 1994, who have been followed up with questionnaires in five stages of their careers, with the last follow-up in 2008. The other sample, the Medical Student Cohort, consisted of all Norwegian medical students who entered medical school in 1993, but no data for this cohort have been used in the present thesis (7-10).

The findings of NORDOC should contribute to the easier identification of individuals who are vulnerable to depressive symptoms and alcohol problems. Some of the results are also assumed to be applicable to other occupational groups. The main goal of this research, of course, was to make life a little better for both physicians and their patients.

1.1 Background

1.1.1 Why study physicians' mental health?

There exists a vast literature on physicians' health. However, the impression it gives is somewhat ambiguous. On the one hand, physicians assess their own health as similar to, or

better than, other employees (2) and have low rates of sick leave (11). The physical functioning of male physicians exceeds that of male academics (12), and the health-related behaviour of female physicians is better than that of other women (13). On the other hand, the results of studies of physicians' mental health show a different trend, and I will present below four reasons why the study of mental-health problems in this occupational group is particularly interesting and important.

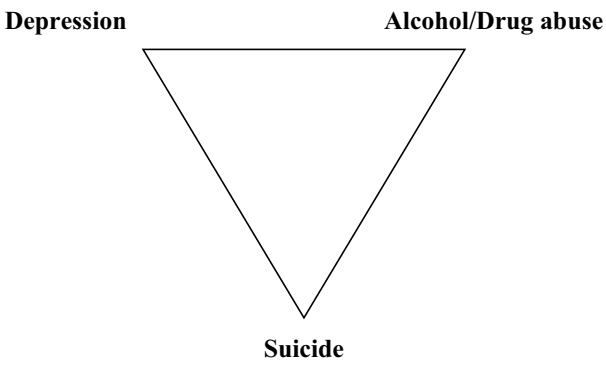
First, representative comparative reports of mental distress and well-being in this population tend to concur. A study of 11,000 British health professionals identified a higher prevalence of minor psychiatric disorders than in the general population (14). A large Norwegian study found that physicians experience more work-related stress than other professionals (15), and another found that male physicians display lower vitality and social functioning than male academics in the general population (12). Finally, one NORDOC study found that physicians report greater dissatisfaction with life than age- and education-matched groups in the general population (16).

Do physicians also experience more mental-health problems, apart from distress and well-being, than comparable groups in the general population? Comparative studies have reported that the depression levels of medical students exceed community norms (17-19), with the internship period specifically associated with elevated levels of depressive symptoms (20;21). Some researchers have described continued elevated depression rates (22-25), estimated at around 30% in the first postgraduate year and gradually falling to below 20%. Conversely, others suggest that levels of depressive symptoms are about the same among physicians and in the general population (18%) (26;27). Several attempts have been made to define the prevalence of depression in physicians beyond their very first years in training, but because of the methodological weaknesses of many of the studies so far undertaken, this remains unclear. For instance, we lack comparative studies based on diagnostic interviews. The early studies of alcohol problems among physicians reported prevalence estimates that were disturbingly high (28;29). However, more recent research has rejected the suggestion that physicians run a greater risk of alcohol problems than socio-economically matched professionals (30-32). Studies of medical students show estimates of hazardous drinking that are similar to (33-35) or even lower than (36) those of other student populations. However, some studies have indicated that there may be one exception to this trend: a persistent finding is that female physicians in their middle and late careers use more alcohol than other women (5;37). Despite this exception, physicians do not seem to have an

increased risk of alcohol problems, and it is somewhat unclear whether their risk of depression is elevated.

The physician is often portrayed as someone who is strong, confers health, provides answers, and offers solutions. This occupational group has a high socio-economic status and is considered by many to be a resourceful, intelligent, and even privileged group in society. Indeed, physicians do enjoy a lower mortality rate from many physical diseases than the general population (38;39). However, the literature unambiguously reports one notable exception: death by suicide (39;40). Male physicians run a modestly (1.2–1.7 times) increased suicide risk compared with males in other professions, whereas the risk of their female counterparts is greatly increased (2–2.7 times) (40-42). Suicide, which increases steeply with age among physicians (41), is strongly linked to depression because depressive disorders are often accompanied by suicidal thoughts (43;44). Several studies have also shown high co-morbidity between depression, suicide, and alcohol problems (45-48). Alcohol problems can lead to additional health problems among those who already suffer from depression (5;6), and the co-morbidity of alcohol problems and depression amplifies an already elevated suicide risk among physicians (40;49). To highlight the relationship between suicide, alcohol problems, and depression, Tyssen has suggested the figure shown below (Figure 1) (9). Data suggest that alcohol problems and depression not only constitute the most common dual diagnosis among physicians receiving psychological treatment (50) but also are two of the most common psychiatric disorders among physicians who commit suicide (51;52).

Figure 1: The SAD triangle (9)



Third, several authors have claimed that physicians are difficult patients because they tend to avoid the patient role, resorting instead to denial and self-neglect. They may be ambivalent about the sick role and find it particularly difficult to acknowledge psychological frailty, in either themselves or their colleagues (53). Many also fear stigmatization (28), and physicians are probably less likely than the general population to seek and receive appropriate treatment for mental disorders (2;54;55). A Finnish study reported that two-thirds of physicians with mental-health problems had treated themselves, and this was the only kind of care they had received (2). The majority of Norwegian physicians who reported mental-health problems requiring treatment had not sought help (55;56). Consequently, self-treatment is common in this occupational group, which leads to delayed recognition and management of their problems (57;58). Interestingly, although physicians have a low rate of sick leave, a disproportionate number of the diseases they report are mental disorders compared with those reported by the general population (2).

Finally, the health and health behaviour of the physician can affect health care in several ways. Not only are physicians role models in the general community with regard to lifestyle and drinking (59) but also physicians' own health practices may affect how they counsel patients (60). We know that physicians fail to detect or treat 40%–60% of cases of depression in their patients (61-63). Furthermore and consistent with this, approximately 40% of individuals who die by suicide had contacted their primary-care physician within a month of their suicide, without eliciting the attention of, or proper action by, the treating physician (6;64). These findings may reflect a corresponding inattention among physicians to their own mental health and to that of their colleagues, which may adversely affect their modelling of self-care and their mentoring of medical students and physicians in training. Finally, an increasing body of research shows that distressed or mentally ill physicians unintentionally put patients at risk of harm (65;66) and that this particularly applies to depressed physicians. A prospective and observational study found that depressed medical residents made medication errors six times more often than non-depressed residents, an effect that, interestingly, was not duplicated among burned-out residents (67). Hence, mental-health problems among physicians are clearly associated with more than the suffering of the individual physician and may be a more important contributor to patient safety than previously suspected. Despite these disquieting findings, it has been noted that "the attitude of the medical profession to the health of its members has always been one of disinterest which is transiently discarded when disaster overtakes one of its members" (68).

A review in the Lancet stated that mental-health problems among physicians constitute an important and under-estimated political health factor because their well-being seems to be an overlooked quality indicator of all health-care systems (69). Communities are entitled to expect a medical system in which the service providers are psychologically healthy, and studies are required to establish risk factors apparent in medical school, to identify individuals at risk and to offer them support and treatment (6).

In summary, although physicians' socio-economic status is high, and their *physical* health seems to exceed that of the general population, the literature expresses concern for the *mental* health of physicians. Despite the uplifting finding that they probably are no more likely to drink hazardously than their contemporaries, young physicians report relatively high levels of depressive symptoms. More established doctors report higher levels of distress and life dissatisfaction than the general population, and they run an indisputably increased risk of suicide. Physicians undoubtedly exhibit suboptimal sickness behaviour, and their mental-health problems may also put their patients at risk. Finally, the culture of medicine may not pay sufficient attention to the topic of mental-health problems among its members. Together, these arguments justify the study of this specific occupational group.

1.1.2 Rationale for this study

The literature on both alcohol problems and depressive symptoms suggests that the onset of both is influenced by a wide range of biopsychosocial factors. Most aetiological models hold diathesis–stress perspectives, in which vulnerability factors account for individual differences in how people respond to the stresses in their lives. These vulnerability factors include biological, personality, interpersonal, and cognitive characteristics (70). This thesis focuses on a selection of these vulnerability factors for two main reasons. First, we aimed to identify risk factors that can be measured in medical school to allow the early identification of individuals at risk (papers II and III). Several studies have shown particularly high levels of distress and depressive symptoms among physicians early in their careers (17;71-73). A growing body of research reports the considerable positive effects of preventive interventions, especially on depression (74;75), which supports our focus on the identification of depressed individuals. Second, because physicians' working conditions vary greatly across the world, research findings on individual factors, as opposed to work-related factors, are presumed to be more widely generalizable.

Each of the longitudinal studies of physicians conducted to date has had weaknesses that reduce the internal and external validity of the findings. Most of these studies have used

availability samples, such as one (38) or two classes (76) from one medical school, or medical officers from only one hospital (22). Consequently, the representativeness of these samples is unclear. Only two studies had an observation period exceeding one year and analysed multivariate models of the predictor variables measured at baseline (56;77). Finally, only one study tested the properties of individual variables for screening mental-health problems requiring treatment in medical school, but it did not identify the individuals at risk with adequate specificity (56).

Two pioneering prospective investigations of physicians were conducted in the USA in the 1970s. Vaillant followed 47 Harvard Medical School students for 30 years and compared their levels of mental-health problems with those of 79 controls (28). He reported in the New England Journal of Medicine that the physicians had considerably higher rates of marital discord and greater use of drugs and psychotherapy than the controls, and that these differences seemed to depend on pre-career factors, such as less-stable childhoods. In the Precursors Study by Thomas and Duszynski, over 1,000 Johns Hopkins' Medical School students were followed throughout a 20-year period (78). That study reported a lack of closeness to parents as one of the most important predictors of later mental-health problems. Both of these studies examined only male physicians, so their findings are not necessarily valid for women. Furthermore, because only respondents from one medical school were investigated, the external validity of the results may be low. Finally, neither of the two studies addressed the early postgraduate years.

More recent studies have focused on the first years of the medical career, but with only short follow-up periods (22;79). When I started to work on this thesis in 2007, only one prospective study of physicians had a follow-up period of more than 3.5 years (80). No recent prospective studies have followed physicians for periods as long as 15 years, so the present study is the first of its type. We had the opportunity to follow the courses of hazardous drinking and depressive symptoms in physicians for an extended period, and we tested the predictive values of individual vulnerability factors for these two important conditions.

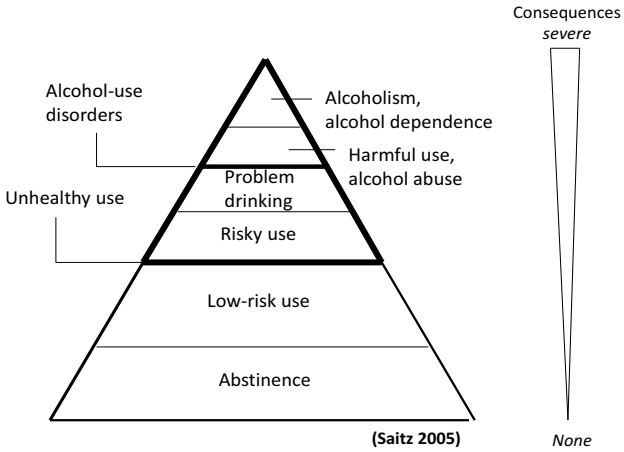
1.1.3 Central concepts

Against this background, we chose to study two central aspects of physicians' mental health: severe depressive symptoms and hazardous drinking. By focusing on these two conditions, we sought to illuminate issues that are especially suited for psychological investigation.

1.1.3.1 Hazardous drinking

Alcohol use can be regarded as a spectrum that extends from abstinence and low-risk use to at-risk use, problem drinking, harmful use, alcohol abuse, and the more severe alcoholism and alcohol dependence (81). This thesis addresses infrequent *hazardous drinking*, which corresponds to *problem drinking* (unhealthy use) (82) in the figure shown below (83). The third item on the Alcohol Use Diagnostic Identification Test (AUDIT), *frequency of drinking to intoxication*, which we used in the present study, has been validated as the best predictor of alcohol misuse in population studies (84).

Figure 2: The spectrum of alcohol use



1.1.3.2 Predictors of hazardous drinking

Expectancy is the subjective probability that a given behaviour will lead to a particular outcome. Social learning theory holds that expectancy exerts a strong influence on behaviour (85). Following this line of argument, expectancy of the effects of alcohol is assumed to predict drinking behaviour (86). Furthermore, the motives for drinking (for example, drinking to cope) constitute the final gateway through which alcohol expectancy is mediated (86). We found these cognitive factors to be particularly relevant to physicians' alcohol use, as demonstrated below.

There is substantial work pressure within the health services (15), and several studies have found that physicians experience considerable work-related stress (20). Although a few

studies have suggested that physicians drink alcohol to cope with stress (28;80) and that the use of alcohol to cope already predicts hazardous drinking in medical school (34;87), the *interplay* between alcohol expectancy and the use of alcohol to cope in physicians has yet to be studied. We are aware of only one long-term follow-up predictor study of alcohol problems among physicians, but that included an all-male sample (the Precursors Study) (88). That study found that the following factors, measured in medical school, were associated with subsequent alcohol abuse: non-Jewish ancestry, lack of religious affiliation, excessive cigarette use, regular use of alcohol, anxiety or anger as a reaction to stress, frequent use of alcohol in non-social settings, past history of alcohol-related difficulties, and maternal mental or alcohol-related problems. No cognitive factors were considered by Thomas and Duszynski.

We do not know whether expectancy of the tension-reducing properties of alcohol can predict hazardous drinking when a long-term window is applied, or whether the possible relationship between alcohol expectancy and hazardous drinking is mediated by the use of alcohol to cope. To the best of our knowledge, only one previous study has explored the interplay between the use of alcohol to cope and alcohol expectancy on hazardous drinking over a period of more than three years (89). However, the sample in that study comprised college students (not medical students), only 29% of whom were men. Therefore, the generalization of these findings to other samples is not straightforward.

1.1.3.3 Depressive symptoms

In this study, depressive symptoms were measured with the *severe depression* subscale, one of the four emotional distress dimensions measured by the General Health Questionnaire-28 (GHQ-28) (90). The subscale is based on seven items, four of which relate to suicidal ideation. We chose this instrument because it has been shown to be an important and valid screening instrument for mental-health problems in the general population (90). Furthermore, it measures symptoms that are associated with suicidal behaviour, which are of particular importance in this occupational group (91). It must also be noted that we aimed to identify depressive *symptoms* rather than diagnostically based depressive disorders. Finally, depressive symptoms must be distinguished from burn-out, although the concepts overlap to some extent. As the definition suggests, burn-out tends to be job related, at least initially, and relatively situation specific, as opposed to depression, which can be generalized across situations and is generally associated with some kind of life stress, such as negative life events (92).

1.1.3.4 Predictors of depressive symptoms

Multicausal models of depression are widely accepted, and much depression research is multidisciplinary. There seems to be a cross-discipline understanding that biological, cognitive, and interpersonal vulnerability factors interact with life stress to produce depressive symptoms (70). Included in these interpersonal factors is dysfunctional parenting. Indeed, some research has particularly pointed to the importance of negative parental experiences for gifted children, such as physicians (93;94). However, much of the research on parenting and depression has been limited by its reliance on cross-sectional or case–control designs (95-97). The two landmark prospective studies of physicians, mentioned above, both found that negative experiences of familial relationships predict depression (28;78). However, neither study controlled for personality, and it is unclear whether the impact of these negative experiences is exerted through personality variables, such as low self-esteem.

A number of empirical studies of non-physician samples have identified associations between suboptimal parenting and depression (97;99;100), as proposed by both cognitive (98) and psychodynamic theories (93). Parker (1993) found that low self-esteem was the personality variable most clearly linked to parental bonding measures, but because such issues cannot be conclusive in single-point cross-sectional studies, he and others (96;101) have recommended that further longitudinal studies be performed to clarify the *pathways* by which deficient parenting can cause vulnerability to adult depression. It is unclear whether perceived suboptimal parental bonding and self-esteem are independent predictors of severe depressive symptoms in physicians, or whether the possible effects of perceived parental bonding are mediated by low self-esteem. This requires clarification in a long-term follow-up study extending beyond 2.5 years, which has only now been undertaken and is reported in this thesis.

Research conducted into so-called “caring professionals” (physicians, psychotherapists, and social workers) indicates that early experiences shape their career choices. Some studies that have compared the early family relationships of caring professionals with those engaged in non-care-related professions have shown that the former are more likely than the latter to report their childhoods as unstable and their parents as distant and overprotective (28;94). Applying psychoanalytic theory, Johnson looked at physicians’ personalities and suggested that a subset of physicians is especially vulnerable to a poor sense of self and low self-esteem as the result of childhood experiences of parental

impotence and emotional neglect. It has been postulated that choosing a medical career may be, in part, an attempt to remedy early emotional neglect by providing an opportunity to give the care and attention to others that one never received as a child (94).

Numerous studies of clinical samples and samples from the general population have found that the experience of deficient parenting plays an important role in the development of depression (102). Despite the disquieting findings of elevated suicide rates and mental distress among physicians described earlier (page 15, section 2), we know of only one relatively recent long-term (10-year) prospective study of physician depression (80), and we lack estimates of the putative risk factors. The majority of studies of physician depression to date have investigated mild and stress-related depressive symptoms (20;103;104), and there have been no studies of the more-severe depressive symptoms that are associated with suicidal ideation, as examined in this thesis (105).

Individual characteristics, such as personality traits and particularly neuroticism, may be important in the development of depressive symptoms (106). Firth-Cozens found self-criticism, a variable that is conceptually strongly related to neuroticism, measured 10 years earlier, predicted depression among general practitioners (80). Similarly, Tyssen and colleagues found that neuroticism predicted both suicidal thoughts (91) and, to some extent, suicidal planning, although these predictions were not retained in the final adjusted models (107). Consistent with these findings are the results of a cross-sectional study that reported self-criticism to be associated with depressive symptoms among male Norwegian physicians (108). In addition to neuroticism, our research group has previously identified deviant personality characteristics ("reality weakness") as possible risk factors for the aggravation of suicidal ideation among physicians (107). However, the predictive validity of these traits requires further investigation in long-term follow-up studies, such as the study reported in this thesis.

Another possible risk factor for depression is female sex, although longitudinal studies of physicians are somewhat inconsistent about whether being female increases the risk of depressive symptoms (38;80). However, studies of Scandinavian medical undergraduates have pointed to an increased risk of distress (109) and depression (106) among female students. This should be further examined in longitudinal studies that include substantial periods in the careers of established doctors.

Some available evidence indicates that there is a peak in depressive symptoms at the beginning of physicians' careers, but there are also cross-sectional reports of elevated rates of depression far into their careers (25;110;111). Therefore, the issue of age as a predictor of

depression among physicians is unclear and requires further investigation in long-term follow-up studies. However, previous depressive symptoms constitute an established risk factor for future depressive symptoms (55;80;91), although risk estimates of such baseline symptoms relative to other risks at baseline are required.

As stated above, a difficult early environment has been found to be associated with depression among physicians (112). Medical-school-specific factors, such as perceived medical school stress (113) and perceived clinical competence and skills (114), may also influence the risk of depression. Previous studies have actually demonstrated a valid measure of medical school stress to be linked to anxiety and depressive symptoms (113). Alcohol problems may also increase the risk of depression, according to the injury hypothesis (115). The factors discussed above should all be included as possible predictors of depression in a representative and prospectively designed study that follows physicians through many years of their established working lives. Representative samples that are large enough to adjust for possible confounding variables are also required. Because depressive episodes may come and go over the years, several observation points will enhance the reliability of the estimates when analysed with repeated-measures statistics.

1.2 Research aims and questions

1) To investigate hazardous drinking in Norwegian physicians, focusing on alcohol expectancy.

a) What is the prevalence of hazardous drinking four and nine years after graduation? (Paper I)

b) Does positive expectancy of the tension-reducing effects of alcohol predict hazardous drinking six years later? (Paper I)

c) Is the possible relationship between alcohol expectancy and hazardous drinking mediated by the use of alcohol to cope? (Paper I)

2) To study depressive symptoms over 15 years in Norwegian physicians and to determine whether parental bonding or other individual factors in medical school predict these symptoms.

a) What is the prevalence of severe depressive symptoms throughout the 15-year postgraduate period? (Papers II and III)

b) Can perceived parental bonding, measured at the end of medical school, predict severe depressive symptoms measured four and nine years after graduation from medical school? (Paper II)

c) Is the possible relationship between parental bonding and depressive symptoms mediated by low general self-esteem? (Paper II)

d) Can individual and stress-related factors measured at the end of medical school (personality traits, hazardous drinking, perceived medical school stress, perceived clinical competence and skills, and parental bonding) predict severe depressive symptoms in the 15 years following graduation from medical school? (Paper III)

3) To examine any sex-based differences in the prevalence of depressive symptoms, hazardous drinking, and the relationships between the variables described above (Papers I, II, and III).

e) Are there sex-based differences in the prevalence of depressive symptoms? (Papers II and III)

f) Are there sex-based differences in the prevalence of hazardous drinking? (Paper I)

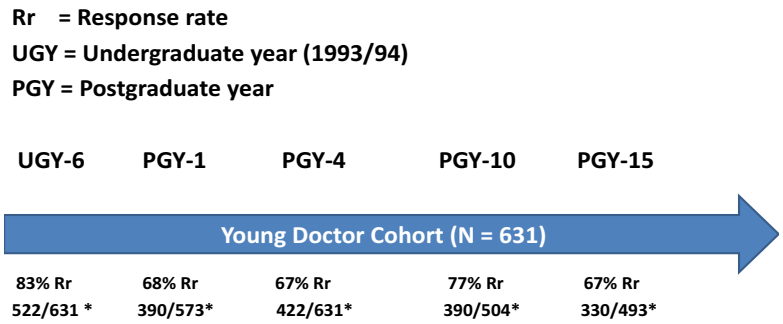
g) Are there sex-based differences in the relationships between the variables studied in papers I, II, and III?

2 MATERIAL AND METHODS

2.1 Study design

This thesis is based on three studies conducted on the Young Doctor Cohort, one of two samples that together constituted the NORDOC study sample. The participants in the Medical Student Cohort (N = 421), which is not part of this thesis, commenced their medical studies in 1993. The participants (of both cohorts) received a comprehensive postal questionnaire (30–40 pages) on five occasions during a 15-year period (1993–2008) (see Figure 3). In paper I, we assessed the capacity of cognitive factors, measured at the third observation point (PGY-4), to predict hazardous drinking six years later. In paper II, we investigated whether the experience of parental bonding reported at the first (UGY-6) and second observation points (PGY-1) predicted severe depressive symptoms measured at the fourth (PGY-10) and fifth observation points (PGY-15). In paper III, we tested the capacity of individual and medical-school-related factors, measured at the first observation point (or at baseline), to predict future severe depressive symptoms, which were measured at all four subsequent observation points.

Figure 3: Study design

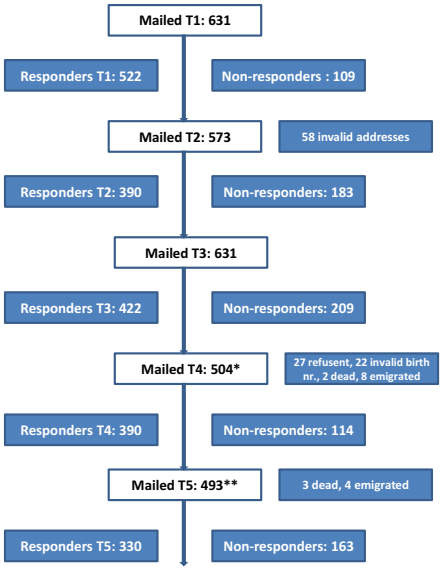


Note: *Number of questionnaires that were mailed.

2.2 Sample

The Young Doctor Cohort consisted of all students who graduated in 1993 and 1994 from all four medical faculties in Norway (N = 631). Eighty-three per cent (522/631) responded at the end of their final term (T1). The mean age at T1 was 29 years, and 57% of the graduates were women. The postal questionnaires were sent out again towards the end of the graduates' medical internships a year later (mean, 1.2 years; SD 0.2) at T2. Seventy-one per cent (371/522) of the T1 sample responded (58% of the original sample; 68% of the questionnaires mailed at T2). Because many of the students had moved after completing medical school, the researchers involved in the project at that time were unsuccessful in tracking the new addresses of 58 (9%) participants (see Figure 3) (55). In their fourth postgraduate year (mean, 3.5 years; SD 0.4) at T3, 81% (422/522) of the baseline sample responded (67% of the original sample; 67% of the questionnaires mailed at T3). The fourth set of questionnaires was sent out during the 10th year after the physicians had graduated (T4), and at this time, 75% (390/522) responded (62% of the original sample; 77% of the questionnaires mailed at T4). The fifth and final set of questionnaires was sent out in the 15th postgraduate year (T5), and 63% (330/522) of the baseline sample responded (52% of the original sample; 67% of the questionnaires mailed at T5). Overall, 42% (219/522) (35% of the original sample) responded at all five observation points, and of these, 53% were women. The mean total observation period was 14.5 (SD 0.6) years, and the average age of all the respondents at T5 was 42 (2.7) years (see Figure 3).

Figure 4: The Young Doctor Cohort



Note: *68 individuals were unaccounted for between T3 and T4; **four individuals were unaccounted for between T4 and T5. Statistics Norway mailed the questionnaires but no longer have the numbers available.

Table 1: Description of the study samples

Level of training/career	Number	Response rate	Women	Paper	Outcome	Design
4 and 10 years after graduation	n = 288	46% (55%*)	55%	I	Hazardous drinking	Prospective
Graduating term, 1, 4, and 10 years after graduation	n = 260	41% (50%*)	55%	II	Severe depressive symptoms	Prospective
All five observation points	n = 522	83% (100%*)^	57%	III	Severe depressive symptoms	Prospective

Note: *% of baseline sample, N = 522; ^GEE analyses required data from at least one of the five observation points for the respondent to be included.

2.3 Dependent variables

2.3.1 Hazardous drinking

"Hazardous drinking" refers to the frequency of drinking to intoxication during the preceding year and was measured by the third item in the original 10-item AUDIT (116). The AUDIT has demonstrated reliability and validity, which compare favourably with, or exceed, those of other much-used alcohol-screening measures (117;118), and it is considered to be one of the best available questionnaires for detecting hazardous drinking in the general population (119;120). Our one-item instrument, AUDIT-3 ("How often do you have five drinks or more on one occasion?"), is one of several short forms of the original AUDIT. AUDIT-3 has been found to identify both male (121;122) and female hazardous drinkers (123) nearly as well as the complete AUDIT. Fleming and colleagues found that the third item was one of two AUDIT items that best distinguished individuals with and without alcohol misuse (according to DSM-III) (84). Our cut-off point, binge drinking at least 2–3 times per month, or so-called "infrequent binge drinking", is in accordance with the frequency of drinking that has been linked to negative health consequences among USA college students (124). In that study, both frequent and infrequent binge drinkers were much more likely than non-binge drinkers to be susceptible to negative psychosocial health problems, such as using no protection during sex, engaging in unplanned sexual activity, getting into trouble with campus police, damaging property, or getting hurt or injured.

Unfortunately, we did not include a measure of lower alcohol consumption, which would have been a more valid measure of hazardous drinking among women (122;123). Hence, it is likely that hazardous drinking among women was under-estimated, and we may have missed predictor effects of this variable among women in paper I, because of type II errors (false negative findings). Although the one-item instrument has been shown to perform well in the detection of hazardous drinkers, as noted above, some authors have questioned its specificity (82). Moreover, a one-item instrument may produce less-reliable results than a multi-item instrument. Nevertheless, because AUDIT-3 has been validated in several other studies, we believe it was also valid in our study (for the male physicians).

2.3.2 Severe depressive symptoms

The GHQ was designed to screen for probable psychiatric disorders in primary-care settings (125), and it has been widely used in research (126). The shorter GHQ-28 is one of several short versions and consists of four emotional distress subscales: somatic symptoms, anxiety/insomnia, social dysfunction, and severe depression (90). We only used the *severe depression* subscale, and examples of the items and scoring procedures used are given in papers II and III. In paper II, we used the mean value of the sum scores at T3 and T4 as a dependent variable. In paper III, we used a dichotomized variable at all five observation points.

The potential reliability and validity of the GHQ-28 have been repeatedly demonstrated and found satisfactory (127;128). Its split-half coefficients and Cronbach's alpha values are generally found to be high. In terms of validity, each test item has been shown to discriminate between respondents who are psychologically distressed and those who are not (129). The Cronbach's alpha values for our sample were 0.82 at T1, 0.98 at T2, 0.89 at T3, 0.91 at T4, and 0.90 at T5. There has been great variation in the cut-off values for the threshold in different settings. Goldberg and colleagues noted that the mean score on the GHQ provides a rough guide to the best threshold (130), whereas others have suggested that the median score should be used (131). Fixed cut-off values have also been suggested, and these have varied from (for GHQ-28) 3/4 to 12/13. In paper III, we applied a threshold of 1, which corresponds to 3/4 with all four subscales included (90). This might be considered a somewhat liberal limit. However, the items that constitute the subscale address symptoms considered to be very severe, as the name implies, and four of the seven items relate to suicidal ideation. Furthermore, we assumed that physicians will under- rather than over-report their symptoms (132), and by applying a relatively low cut-off, we hoped to

identify most individuals with clinically relevant symptoms. Finally, we tested the internal validity of our cut-off by calculating the responses that overlapped other clinically relevant categorical variables. For details, see paper III, pages 5–6. We found that this internal validation supported our decision to use 1 as the cut-off value.

A preliminary analysis revealed relatively low stability rates for severe depressive symptoms (see paper III, page 7). This is consistent with reports of low test–retest correlations for the GHQ compared with other instruments that screen for depressive symptoms and is probably attributable to the exact wording of the items. The response categories are “less than usual” and “more than usual” rather than a more objective evaluation of symptoms. This leaves the definition of “usual” to the respondent, and clinically depressed individuals may rightfully answer “no more than usual” when they feel that life isn’t worth living, whereas a slight decline in mood may lead usually happy individuals to respond “more than usual” to questions about their symptoms. Whatever the objective levels of their symptoms, the responses of most people will vary over time, alternating between “more” and “less” than usual. However, it seems likely that the most-depressed individuals will tend to answer “more than usual”. We wanted to identify a reliable measure of this tendency, so we included the depressive symptoms measured at four observation points. However, the relatively low stability of these symptoms may also reflect the variation in the expression of symptoms of affective disorders with the years (70).

Table 2: Dependent variables, with points of measurement

	T1 Graduating semester medical school	**T2 1st postgraduate year**	**T3 4th postgraduate year**	**T4 10th postgraduate year**	**T5 15th postgraduate year**
Paper I					
Hazardous drinking	–	–	–	16% (M) 2% (F)	–
Paper II					
Severe depressive symptoms (mean score)			8.0 (SD 2.3)	8.2 (SD 2.7)	–
Paper III					
Severe depressive symptoms		7.7%	6.9%	10.4%	7.2%

2.4 Independent variables

2.4.1 Demographic characteristics

Age was measured as a continuous variable. Sex was coded 1 for women and 2 for men. We tested the curvilinear effects of age on both hazardous drinking and severe depressive symptoms but found none.

2.4.2 Alcohol expectancy

The expectancy that alcohol reduces tension was measured with the Tension Reduction Subscale of the Alcohol Expectancy Questionnaire (TR-AEQ) (133). The TR-AEQ is a nine-item self-report measure that assesses each respondent's expectation that a moderate dose of alcohol will result in less anxiety or a reduction in tension. See paper I (page 19) for the scoring details and examples of items. The internal consistency in our sample, as measured by Cronbach's alpha, was 0.83, and the score distribution was close to normal.

Some researchers have questioned the discriminative ability of the AEQ subscales and have suggested that in reality, the different scales tap a global positive attitude towards the effects of alcohol (134). Consequently, caution should be exercised when interpreting findings that incorporate one or more of these subscales as specific to that particular subscale (135). Another critique of alcohol expectancy research, in general, is that the

expectancies explain only a small proportion of the variance. Hence, these constructs cannot constitute a sufficient theory of alcohol problems (136).

2.4.3 Use of alcohol to cope with tension

This was measured with one question, which was also used in a national survey in the United States (87;137): “When you feel worried, tense, or nervous, do you ever drink alcoholic beverages to help you handle things?” This has previously been validated in the two cohorts of NORDOC at the undergraduate level (34;87). The respondents were given four alternatives, but the responses were dichotomized as “never” and “with any frequency” because we considered the important distinction to be whether this type of coping was used or not (34;87). In Norwegian culture, such drinking behaviour is so stigmatized, at least among medical students, that the only clearly negative response would be “never”. Because the use of alcohol to cope was measured with only one item, its reliability may be low, and the conclusions drawn in paper I should be interpreted somewhat cautiously.

2.4.4 Hazardous drinking

When hazardous drinking was used as an independent variable, it was assessed somewhat differently than in paper I. In paper III we used a slightly modified version of the AUDIT (116). One of the 10 items (item 2: “How many drinks containing alcohol do you have on a typical day when you are drinking?”) was omitted from the questionnaire based on the rationale that this item may be misunderstood in the Norwegian context. “A typical day when you are drinking” may be interpreted as a typical day with a party or any other day when you are drinking alcohol. The version used here has previously been used to study hazardous drinking among Norwegian doctors, police, and ambulance officers (138).

To calculate readily comprehensible and meaningful risk estimates (odds ratios), we dichotomized the continuous predictor variables in paper III, despite the potential loss of information entailed by this strategy. Consequently, the median and above-median scores for hazardous drinking were grouped as “high”, and scores below the median as “low”. Although this is a common statistical strategy for comparing high and low scorers in a sample, it can be claimed that the median and above-median scorers do not necessarily drink hazardously in this particular sample. The score distribution was rather skewed, and the median value of 3 was relatively low (range 0–40). There has been dissent about the most

appropriate cut-off for AUDIT (139-141), and a cut-off lower than 5 has so far not been considered to differentiate individuals without potential problems from those with them.

Other studies have also suggested that self-reported alcohol consumption tends to be under-estimated (142). Investigations have shown that the amount of alcohol sold is double the amount of alcohol reportedly consumed (143). Medical students have been reported to fear stigmatization, and as representatives of the health system (132), they may be even more inclined to under-estimate their alcohol consumption, which lends support to the rather low cut-off value used. Therefore, we assume that we identified the drinkers potentially at risk.

2.4.5 Parental bonding

The Parental Bonding Instrument (PBI) is a 25-item inventory that asks respondents to rate the attitudes and behaviours of their parents during their first 16 years of life in terms of two dimensions: care and overprotection (144). The PBI has gained wide acceptance as a robust measure of perceived parenting and is the most consistently used measure of perceived parenting in clinical and non-clinical subject groups (145). See paper II, page 23, for details. The psychometric properties of the PBI have been extensively examined and found to be satisfactory (144;146;147). The Cronbach's alpha values in the present sample were 0.91 (paternal care), 0.89 (paternal overprotection), 0.92 (maternal care), and 0.91 (maternal overprotection).

As mentioned above, the PBI separates parenting into the two components of care and overprotection, with optimal parenting comprising perceived high care and low overprotection. It has previously been suggested that parents perceived to be less caring and more overprotective, conceptualized as "affectionless control", were characteristic of subjects with adult depression. However, more recent work has given primacy to reported lack of care and has attached less importance to the overprotection dimension (148;149). Therefore, we chose to study the dimensions separately, contrary to the traditional approach of creating categories of high/low care and high/low overprotection (150). Furthermore, because we demonstrated in paper I that a perceived lack of maternal care had the strongest effect on depressive symptoms, we only included this dimension in the analysis of paper III.

The PBI was randomly applied to half the sample at T1 and to the remaining half at T2. This was done to limit the size of the comprehensive questionnaires (30–40 pages in total) and because these perceptions of parental bonding were presumed to be consistent

over time (151-153). When these two subsamples (T1 and T2) were compared for any differences in the mean PBI scores using *t* tests, no significant differences were found.

Because the PBI measures parental behaviours retrospectively, it is open to the possible influence of current moods or recall bias. However, the recollection of one's parental environment is generally substantially unaffected by sex, depression history, or life experience, and subjects' perceptions do not change with fluctuations in depression or neuroticism levels (151;153;154). Some studies have also shown that perceived parenting, as measured with the PBI, is related to actual parenting (155;156).

2.4.6 Self-esteem

General Self-Esteem (GSE) was measured at T1 by eight items from the neuroticism scale in Torgersen's Basic Character Inventory (BCI). The original BCI neuroticism scale suggested by Torgersen has a total of 18 items, and the GSE and BCI neuroticism nine-item scales (see description of BCI below) overlap on only two items: "My lack of self-esteem is sometimes a bother for me" and "Sometimes I think that everybody else does things much better than I do". Each item of the GSE was measured on a four-point scale ranging from 1 (little agreement) to 4 (strong agreement), with higher summed scores indicating higher general self-esteem (range 8–32). Cronbach's alpha was 0.88.

The test–retest reliability for GSE measured at T1 and T2 was 0.70, which confirms the high stability of GSE, as expected, indicating that this is a trait rather than a measure of state. Because GSE is strongly related to BCI neuroticism, this high stability also supports a high potential reliability of the data derived from this personality inventory, at least with regard to the neuroticism scale.

2.4.7 Personality traits

These were measured in a randomly selected half of the sample in the graduating semester (T1) and in the remaining half of the sample in the year after graduation (T2), using the 36-item version of Torgersen's BCI (157;158).This is a modified version of a questionnaire, grounded in psychodynamic personality theory, that Lazare, Klerman, and Armor constructed in 1966 (159). Torgersen has reported considerable heredity for the three main dimensions, corresponding to Eysenck's Giant Three personality dimensions: neuroticism (vulnerability), extraversion (intensity), and conscientiousness (control, constraints) or low psychoticism (157). A fourth dimension, called "reality weakness", was added by Torgersen

(158). Reality weakness includes perceptions and ideations on the borderline between reality and fantasy. It measures chronic illusions, paranoid traits, and problems with identity insecurity and relationships, traits that are associated with severe personality disorders. Examples of items that measure this trait are: “It is difficult for me to trust people because they so often turn against me or leave me in the lurch”, “Sometimes I feel like I live in a fog”, and “Every now and then, I get strange thoughts in my head that I can’t get rid of”. Each dimension is measured as a dichotomous response (agree/do not agree) on nine items, thus initially giving a continuous variable ranging from 0 (low) to 9 (high) per dimension. However, like the other continuous independent variables examined in paper III, the personality dimensions were divided into groups of “high” (median or above-median) and “low” (below-median) scores.

The Cronbach’s alpha values for the scales were 0.82 for neuroticism, 0.87 for extraversion, 0.73 for conscientiousness, and 0.65 for reality weakness. Thus, the neuroticism and extraversion dimensions showed the greatest reliability, whereas the reliabilities of both conscientiousness and, in particular, reality weakness were somewhat lower. This reduces the precision of the measurements, which in turn increases the risk of type II errors, weakening the validity of our findings. The BCI has been validated in both clinical and non-clinical studies (157;160-164), and it has been used previously to assess Norwegian medical students and physicians (55;56;165;166). The subscales also resemble the personality inventories used in other studies, like those described by Eysenck (167), Cloninger (168), and Costa and McCrae (169).

However, it is possible that the BCI does not capture all aspects of personality, only the Big Three. Other inventories, such as the Revised NEO Personality Inventory (NEO PI-R) (170), which also assesses agreeableness and openness to new experience, probably evaluate a larger domain of personality traits than the BCI.

2.4.8 Perceived medical school stress

This was measured with a slightly modified version of Vitaliano’s instrument (113). The PMSS includes several areas of medical school stress, such as academic demands, endurance, and limited recreational time. Bramness et al. (3) translated Vitaliano’s instrument into Norwegian, omitting one item about electives and clerkships that was regarded as irrelevant to Norwegian students, and adding one item about stress concerning accommodation. Examples of the items are “Medical training controls my life and leaves too little time for other activities” and “Medical school is more of a threat than a challenge”.

The respondents reported on a five-point scale from one (strongly disagree) to five (strongly agree), and the summed scores for the 13 items were used to indicate a high/low level of stress (range 13–65) However, like the other continuous independent variables assessed in paper III, the PMSS scores were divided into groups of "high" (median or above-median) and "low" (below-median) scores. Cronbach's alpha was 0.78. Studies in which this instrument has been used have reported associations between perceived medical school stress and anxiety and depressive symptoms (3;113). The predictive validity of PMSS in assessing mental-health treatment needs has previously been documented in both medical undergraduates (171) and postgraduates (56), but its predictive validity beyond the fourth postgraduate year has never before been investigated.

2.4.9 Perceived diagnostic skills

This variable was measured at T1 with an instrument constructed specifically for use in this study. The 16 items used to assess this variable were derived from a factor analysis (principal components, varimax rotation) of 32 original items concerning physicians' perceived skills in medical diagnosis, their perceived recording skills (see below), and their identification with the physician role. The perceived diagnostic skills variable describes how certain/uncertain the respondent feels about the way to approach a case history and a physical examination based on a specified medical diagnosis. Examples of the items used to assess this variable are "To what degree do you have the technical and theoretical foundation to evaluate the findings consistent with", followed by: "a diagnosis of cancer of the prostate"; "finding an enlarged liver", etc. The respondent was given seven response alternatives from no foundation (1) to a good foundation (7), and a high summary score for the 16 items indicated a high level of perceived skills (range, 16–112). However, like the other continuous independent variables in paper III, the perceived diagnostic skills variable was divided into groups of "high" (median or above-median) and "low" (below-median) scores. Cronbach's alpha was 0.84.

2.4.10 Perceived recording skills

As mentioned in the paragraph above, the six items that constituted the perceived recording skills variable were derived from a factor analysis of 32 original items. This variable was measured by six items based on previous patient interviews and medical records. The items measured the respondent's confidence in his/her work during previous patient interviews

and when medical records were written up during the previous term. Examples of the items are: “I manage to direct the interview so that I get the relevant information about the past medical history” and “I easily get an overview of the progress of the present illness”. The response alternatives ranged from low (1) to high (7), and the scores were added to a sum score (range, 6–42). Like the other continuous independent variables in paper III, the perceived recording skills variable was divided into groups of “high” (median or above-median) and “low” (below-median) scores. Cronbach’s alpha was 0.77.

A critique relevant to both the perceived diagnostic and recording skills (described above) is the fairly weak concordance reported between the self-assessment of clinical skills and expert ratings (172;173). Generally, females tend to assess their own skills less favourably than males (174;175). Consequently, these variables may reflect self-esteem rather than actual skills. Ideally, we should have objective measures of these traits. However, perhaps *perceived* rather than *objective* competence is more important in the development of depressive symptoms. Furthermore, when we control for neuroticism in the adjusted predictor model, we also adjust for the tendency to self-report negatively (176).

Table 3: Description of responses to the independent variables at each observation point

	T1	T2	T3	T4	T5
All papers					
Age	28 (SD 2.8)	29 (2.8)	32 (2.8)	38(2.6)	43 (2.6)
Sex	57% women	56%	56%	59%	58%
Paper I					
Alcohol expectancy	–	–	2.0 (0.6)	–	–
Alcohol to cope	–	–	8%	–	–
Paper II					
Maternal care*	28.5 (6.5)	–	–	–	–
Maternal overprotection	10.7 (7.0)	–	–	–	–
Paternal care	25.4 (7.0)	–	–	–	–
Paternal overprotection	9.1 (6.2)	–	–	–	–
Self-esteem	17.0 (3.7)	–	–	–	–
Paper III					
Neuroticism*	3.49 (2.31)	–	–	–	–
Control	3.05 (2.03)	–	–	–	–
Extraversion	5.56 (2.46)				
Reality weakness	1.09 (1.42)	–	–	–	–
PMSS	20.35 (6.85)	–	–	–	–
Recording skills	28.64 (4.75)	–	–	–	–
Diagnostic skills	85.57 (10.24)	–	–	–	–
Hazardous drinking	–	–	9%	8%	–

Note: *measured in a randomly selected half of the sample at T1, and the other half at T2.

2.5 Statistical methods

In papers I and II, the statistical analyses were performed with the statistical package SPSS version 17. In paper III, we used PASW 18. The level of significance was set at $p < 0.05$ for all papers.

Paper I: Mean differences between the scores of men and women on the TR-AEQ were tested with *t* tests, and differences in hazardous drinking and the use of alcohol to cope with tension were tested with odds ratios (ORs). A logistic regression analysis was performed to determine whether alcohol expectancy scores at T3 predicted hazardous drinking at T4 (Table 3). We included the predictor variables in blocks to test the effects of separate variables and to identify any indication of mediation. See paper I, page 20.

Paper II: The distributions of severe depressive symptoms at T3 and T4 were considerably skewed, so the severe depressive variable was log (ln) transformed twice, which resulted in fairly normally distributed standardized residuals. *t* tests were used to test for sex-based differences in the mean levels of reported severe depressive symptoms and parental bonding dimensions. The effect of perceived parental bonding on severe depressive symptoms and the possible mediating role of GSE were tested with a linear multiple regression model and subsequent path analyses. We chose the method suggested by Baron and Kenny (177), in which three conditions must be fulfilled: 1) the independent variable (maternal care) should be significantly related to both the dependent variable (depressive symptoms) and the potential mediator (self-esteem); 2) the beta value of the independent variable should be considerably reduced when the potential mediator is included; 3) the beta value of the potential mediator should not be markedly reduced when the independent variable is included. Although reduced, the beta value for maternal care was still significant when we controlled for self-esteem, so the mediation was partial rather than complete.

Paper III: The effects of the predictor variables on severe depressive symptoms at T2–T5 were tested with generalized estimating equations (GEEs) (178). GEEs are an extension of the generalized linear model, which accounts for correlated repeated categorical measures within subjects. Despite the loss of variance brought about by the dichotomization of variables, we dichotomized the continuous predictor variables (except age) and the outcome variable, severe depressive symptoms. This was done because an important aim of this study was to calculate easily comprehensible and meaningful risk estimates (ORs). The predictor variables were divided into “high” (median or above-median) and “low” (below-median) scores.

We expected severe depressive symptoms at T1 to be a strong predictor of future severe depressive symptoms, and two predictor models were applied (GEEs do not allow block-wise analysis). The first multivariate model included all the significant univariate

predictors, except severe depressive symptoms at T1, whereas the severe depressive symptoms at T1 variable was added into the second model. See paper III for details about the estimation of the differences in prevalence and the interaction effects. As regards the strength of the predictor effects, there exists no estimate corresponding to the explained variance (R^2) in GEEs.

GEEs provide a powerful tool for analysing longitudinal data. They model time effects by properly accounting for correlations between repeated measurements, and they handle missing data more appropriately than traditional models (179). However, the validity of the data rests on the assumption that missing data are missing completely at random (MCAR), and systematic patterns in attrition can give biased results (180). We ran *Little's test for MCAR*, which showed that the data were not MCAR. Therefore, we compared the subjects who did not have valid information regarding depressive symptoms at T2 with those without missing data. The two groups were not substantially different in terms of either their hazardous drinking or their depressive status at T1. This does not exclude the possibility of bias, but it provides some reassurance that the subjects with incomplete data resembled the remaining subjects, at least in terms of the variables we considered to be most relevant: hazardous drinking and severe depressive symptoms.

2.6 Methodological issues

2.6.1 Study design

This study had a longitudinal design, with five observation points. Only one year separated the first two points, whereas there were three years between the second and third observation points, and five years between the third and fourth and between the fourth and fifth observation points. This longitudinal design allowed us to explore the courses of our outcomes of interest, and to investigate possible predictors of these. However, we may have missed important information, and this is particularly relevant to the last observation points because the time windows were quite long. Variables beyond our focus might also have affected the relationships between our predictors and outcomes. One- or two-year time spans, or even shorter, have been reported to be optimal in the prediction of mental health, particularly regarding work-related stress (176;181;182). Nevertheless, our main aim was to investigate whether we can already identify in medical school the possible risk factors for the development of depressive symptoms many years later.

2.6.2 Psychometric considerations

2.6.2.1 Reliability

"Reliability" refers to the replicability of our measurements, or our ability to obtain the same answer across different tests, typically items in a test battery (internal consistency), across different testers (inter-rater reliability), and across different time periods (test–retest reliability). The most relevant form of reliability in the present discussion is internal consistency. Cronbach's alpha is a measure of internal consistency based on the average intercorrelations within a set of items presumed to tap the same underlying phenomenon (dimension). This coefficient then estimates the proportion of the variance on an additive scale based on the pool of items that is attributable to the common underlying factor. Coefficients above 0.70 thus indicate that more than 70% of the variance in an index is accounted for by the underlying ("latent") phenomenon and is, by convention, generally considered good enough. With the exception of a Cronbach's alpha value of 0.65 for the personality dimension *reality weakness*, all of the instruments we used showed acceptable reliability according to this criterion. Whereas reliability says something about precision, validity is a measure of the accuracy of an instrument; i.e., the extent to which the instrument actually measures what it is supposed to measure. Because reliability is a necessary precondition for validity, the high Cronbach's alpha values obtained for our measurements are reassuring. Nevertheless, reliability is no guarantee that validity has actually been achieved.

2.6.2.2 Validity

As mentioned above, validity is the extent to which an instrument measures what it is supposed to measure, and it is commonly divided into three broad categories. *Content validity* refers to the apparent relevance of the instrument items; *criterion validity* refers to the instrument's correlation or overlap with other measures of the same phenomenon, either at present (concurrent validity), in the future (predictive validity), or in the past (retrospective validity). We tested the predictive validity of several variables in our study. For instance, in paper I, we tested whether alcohol expectancy predicted hazardous drinking five years later.

Construct validity is related to the theoretical ideas behind the phenomenon under consideration and can be viewed as the overarching, perhaps most important, form of validity. This involves testing the degree to which the instrument correlates strongly with

theoretically related measures (convergent validity) and weakly with unrelated measures (divergent validity); e.g., whether our measure of severe depressive symptoms is or is not a valid measure of the construct *depressive disorder*. The GHQ *severe depression* subscale focuses on the most serious (i.e. suicidal) symptoms, so we may have missed milder depressive symptoms (sadness, fatigue, tension, etc.). However, because physicians run an increased risk of suicide, we regarded the use of this instrument as particularly relevant to this population. Validation reports of our measure *hazardous drinking* have primarily studied students and young adults, so its construct validity for a middle-aged population is weaker.

2.6.2.2.1 Selection bias

How does sample attrition influence results? Are distressed physicians less likely to complete surveys because of apathy or are they more likely to complete surveys related to their jobs and mental-health issues because of greater interest ("saliency")? There is conflicting evidence regarding the possible biasing effect of survey non-response. A broad population survey found that people who self-assessed their health as poor had a slight tendency to fail to co-operate in surveys (183). One study of physicians found that estimates made from early responders closely approximated those obtained at the conclusion of a long field period (184). Mykletun and colleagues found that people with mental-health problems participated in a longitudinal health survey to a lesser extent than healthy subjects (185). However, they concluded that this had a stronger influence on prevalence estimates than on estimates of the associations between variables. Surveys concerned with alcohol or substance use have produced mixed results. One investigation reported no systematic relationship between the probability of response and the amount of alcohol consumed (186). On the contrary, two other studies (187;188) found that heavy alcohol users were more likely than light users to be non-respondents.

Attrition analyses have shown no significant associations between severe depressive symptoms, sex, or age and the number of times that subjects participated in surveys. Our sample of responders was not biased with respect to previous severe depressive symptoms, sex, or age. Furthermore, we found no significant difference in the occurrence of hazardous drinkers at T1 among those who responded at T2 and those who did not. However, in terms of personality, the non-responders at T3 had somewhat lower levels of extraversion than the responders, which has been reported previously (56). This is consistent with the suggestion

that extraverted individuals are more willing to self-disclose. However, we consider this unlikely to influence our conclusions.

These findings notwithstanding, we still consider it more likely that physicians who struggle most in their lives, be it with alcohol problems or depressive symptoms, are less likely to tackle a 30–40-page questionnaire, regardless of the relevance of the issues covered. This implies that our prevalence estimates are lower than the true prevalences and that we run the risk of false negative findings (type II errors).

2.6.2.2.2 Reporting bias

The diagnosis of depressive symptoms was based on self-report and was not confirmed by an inventory designed to identify depressive disorders or a structured clinical interview. Therefore, we should refer to severe depressive symptoms rather than depressive disorders. However, data that have been sampled by surveys are not contaminated by the encounter with an interviewer, or by factors such as age and gender differences between the interviewer and respondent. The results of experiments tend to indicate that a high degree of anonymity in the test situation reduces the scores on impression management scales (189-191); i.e., reduces the subject's tendency to respond in a socially desirable manner. Still, impression management may influence the respondent even when the interviewer is not present. Some researchers have shown that, in at least some study populations, heavier drinkers and drinkers with problems tend to under-report their drinking to a greater extent that lighter drinkers (192-194). Therefore, we could also risk underestimating the number of hazardous drinkers in our sample. However, other studies have reported that bias is minimal in surveys of substance abuse as long as anonymity is assured (195). Reviews of the validity of self-reported alcohol consumption have drawn conflicting conclusions (142;192;196;197).

One study found that the perception of the stigma associated with mental illness was prevalent among medical students throughout the medical profession, and hampered appropriate help-seeking behaviour (132). Therefore, we can infer that experiencing a mental-health problem may also have been viewed as a form of weakness by the physicians in our study sample, causing them to under-report such problems in their responses. In conclusion, we assume that the validity of our results is somewhat weakened by reporting bias.

2.6.2.2.3 External validity

Sufficiently high response rates are a prerequisite for external validity, which is the degree to which the results can be generalized beyond the current sample. The response rates in the present thesis were better than those often reported in longitudinal and other studies of physicians (198). In fact, low response rates to surveys are common among physicians (199;200). A study that examined the response rates to a large number of mail surveys published in medical journals showed that non-physician-directed surveys had a mean response rate of 68% compared with 54% to physician-directed surveys (201). It is our general impression that the inclusion of many work-related variables, such as work stress, at T4 and T5 inspired physicians to respond. There is some evidence that physicians tend to respond more often to questionnaires that they find highly relevant (202). As many as 75% and 63% of the baseline sample responded at the 10- and 15-year follow-ups, respectively. Hence, the response rates obtained were remarkably high, taking into account that this was a postal questionnaire of 30–40 pages in a 10- and 15-year follow-up of a nationwide sample (203). By comparison, cross-sectional studies in North America seldom report response rates above 50%, and some are even considerably lower (204;205).

The working conditions for physicians differ considerably across countries. Work hours are highly regulated among Scandinavian physicians compared to that in many other countries. This is caused by strong unions for physicians and a relatively socialized health care system. Are the findings of this thesis therefore generalizable to physicians working in other countries? This was precisely one of the reasons for choosing individual vulnerability factors; we wanted to investigate factors that were assumed to apply regardless of hours worked per week, etc. Consequently, we consider the findings that have emerged from our work to have external validity.

However, there are also common features in the working and living conditions across Nordic countries, and we believe that the findings of the NORDOC studies are generalizable to these countries, even when working conditions have been included.

2.7 Ethics

To ensure the respondents' confidentiality, the data file identities were matched with the participant identities by Statistics Norway. Therefore, only anonymous data were presented to the researchers. The study was conducted according to the guidelines of the Regional Committee for Medical Research Ethics in Norway with the approval of the Norwegian Data Inspectorate.

3 SUMMARY OF RESULTS

3.1 Paper I

Alcohol expectancy and hazardous drinking: a 6-year longitudinal and nationwide study of medical doctors

Objectives: The study's aim was to test a tenet of social learning theory: to determine whether physicians' expectancy that alcohol use reduces tension predicts the extent of their hazardous drinking, and whether this effect is mediated by drinking to cope with stress.

Findings: There was a sex-based difference in hazardous drinking, with more male hazardous drinkers at both measurement points. At T1, 14.5% of the men and 3.7% of the women reported hazardous drinking. At T2, these percentages were 16% and 2%, respectively. The changes in the levels from T1 to T2 were not significant. Sixty-five per cent (13/20) of the men were hazardous drinkers at both observational points (OR = 21.0; CI= 6.7–66.1, $p < 0.001$). The men reported a higher expectancy that alcohol use reduces tension than did the women ($p < 0.01$), and there was an almost significant sex-based difference in drinking to cope (12% of men and 6% of women, $p = 0.07$). Predictors of hazardous drinking at T2 were male sex (OR = 5.8; CI = 1.6–20.5, $p < 0.01$), alcohol expectancy (OR = 3.4; CI = 1.4–8.6, $p < 0.01$), and hazardous drinking at T1 (OR = 14.7; CI = 4.8–45.2, $p < 0.001$). The effect of alcohol expectancy on hazardous drinking at T2 was not mediated by drinking to cope. Hazardous drinking at T1 mediated the effect of drinking to cope on hazardous drinking at T2.

Conclusion: Efforts to reduce drinking among medical students and doctors should target both alcohol expectancy (beliefs) and hazardous drinking (behaviour). Our measure of hazardous drinking may have under-estimated hazardous drinking among women, and our findings are most valid for men.

3.2 Paper II

Parental bonding and self-esteem as predictors of severe depressive symptoms: a 10-year follow-up study of Norwegian physicians

Objectives: Elevated rates of suicide and depression among physicians have been reported. The association between perceived parental bonding and depressive symptoms has yet to be studied longitudinally in this occupational group. We examined the perception of suboptimal parental bonding as a predictor of severe depressive symptoms in a nationwide cohort and evaluated whether low self-esteem mediates this relationship. After graduation (T1), medical students (N = 631) were followed up for one (T2), four (T3), and 10 (T4) years.

Findings: There were no sex-based differences in the mean depressive scores. The mean depressive scores at T3 were 8.0 (SD 2.3) and 8.2 (2.7) at T4 (range 7–28). Female physicians reported higher levels of care from their mothers ($p < 0.05$) and less overprotection from their fathers ($p < 0.05$). Low care from the mother predicted severe depressive symptoms (beta = 0.003; CI = 0.01–0.001; $p < 0.05$), an effect that was stronger for male physicians than for female physicians (there was a significant interaction between low mother care and male sex). The relationship between perceived parental bonding and depressive symptoms was partially mediated by low self-esteem in both sexes.

Conclusion: Teachers involved in medical training should be aware that negative perceived parental bonding, sometimes expressing its effect through low self-esteem, is a risk factor for severe depressive symptoms. Efforts should be made to identify and subsequently to offer help to medical students at risk, and particular attention should be paid to low self-esteem in these students.

3.3 Paper III

Risk factors at medical school for later severe depression: a 15-year longitudinal, nationwide study (NORDOC)*

*The Longitudinal Study of Norwegian Medical Students and Doctors

Objective: Physicians have an increased risk of suicide compared with academics and other health professionals, and depressive symptoms are common among young physicians. We lack prospective studies of physician depression that identify the early risk factors. The aim of this study was to investigate personality traits and other possible risk factors for severe depressive symptoms over an extended period of physicians' medical careers (15 years).

Findings: Overall, there was a significant reduction in the levels of reported severe depressive symptoms from T1 to T2 ($p < 0.01$) and from T1 to T5 ($p < 0.01$). At T1, 13.7% of the physicians reported severe depressive symptoms; this percentage was 7.7% at T2, 6.9% at T3, 10.4% at T4, and 7.2% at T5. There were no sex-based differences in these rates. The following variables were independent risk factors for future depression: high neuroticism (OR = 3.40, CI = 1.53–7.57; $p < 0.01$); high reality weakness (OR = 2.28; CI = 1.24–4.21; $p = 0.01$); young age (OR = 1.12; CI = 1.04–1.21; $p < 0.01$); and severe depressive symptoms at T1 (OR = 3.62; CI = 2.13–6.14; $p < 0.001$).

Conclusion: As well as young age, severe depressive symptoms at T1 and high neuroticism predicted a 3–4-fold increased risk of severe depressive symptoms over the 15-year follow-up period, whereas high reality weakness doubled the risk. These factors should be targeted in medical school.

3.4 Additional preliminary findings not published in the papers

Alcohol problems and depression:
A 15-year longitudinal, nationwide study of Norwegian physicians

Objective: Extensive research over the last four decades has attempted to establish the relationship between alcohol use and depression, but the results of longitudinal studies have been conflicting. Longitudinal prospective studies with several observation points are required to address properly the directionality and possibly causal nature of the association between the two conditions (206). Within the medical profession, the relationship between

alcohol problems and depressive symptoms has only been studied among undergraduates (207). As far as we know, the present study is the first to address these issues in a representative sample of more-established working physicians. We investigated whether alcohol problems are more likely to precede or follow depressive symptoms. The first option supports the so-called *injury hypothesis* (115), and the second alternative supports the *self-medication hypothesis* (208).

Findings: First, we tested the predictive effect of severe depressive symptoms (GHQ-28) on alcohol problems (AUDIT) using linear repeated measures. Three time periods were modelled prospectively using a linear repeated-measures analysis. To make the tests prospective (lagged), we used the independent variable measured at the preceding time point. Hence, depressive symptoms measured at T1, T2, and T3 were paired with alcohol problems measured at T2, T3, and T4, respectively. In the corresponding manner, alcohol problems at T1, T2, and T3 were paired with severe depressive symptoms at T2, T3, and T4, respectively.

We failed to find that severe depressive symptoms, measured at the preceding time point, predicted alcohol problems ($p = 0.166$). Nor did we show that alcohol problems significantly predicted severe depressive symptoms ($p = 0.66$). Therefore, we divided the AUDIT items into the three subscales that are considered to make up the AUDIT (209): 1) alcohol consumption, 2) harm, and 3) dependence. Linear repeated models with depressive symptoms as the dependent variable found that one of the three subscales, alcohol dependence measured at the preceding time point, predicted severe depressive symptoms (fixed effect: beta = 0.013, $p = 0.02$).

Conclusion: Our finding that drinking (AUDIT dependence) precedes depressive symptoms supports the *injury hypothesis*, rather than the *self-medication hypothesis*.

4 DISCUSSION OF THE RESULTS

4.1 Hazardous drinking in Norwegian physicians, with a focus on alcohol expectancy

4.1.1 Prevalence of hazardous drinking

The prevalence of hazardous drinking was 8% four years after graduation and 9% nine years after graduation. The most valid estimates were obtained for the male physicians: 15% and 16%, respectively. As reported in an earlier publication describing the NORDOC data, the prevalence of hazardous drinking in the final semester of medical school was 12% (22% for men and 5% for women) (87). The prevalence estimates remained stable between the last year of medical school and the ninth postgraduate year, both when the sexes were considered together and separately.

The ways of measuring hazardous drinking used differ across studies, which makes any exact comparisons impossible. However, our prevalence data collected in medical school seem somewhat lower than those reported in final-year medical students in Britain (33) and the USA (5), whereas they correspond well to recent Swedish data (36). Prevalence estimates of hazardous drinking among physicians in Finland seem to exceed ours (210), as do those reported for German (211) and Australian (212) physicians. Prevalence estimates reported from the USA have been both higher and lower (213) than those reported in the present study. However, as mentioned earlier, we lacked a sex-specific question on hazardous drinking among women, and this may have produced an artificially low estimate.

Norway, like Sweden, is a country that is relatively temperate in terms of alcohol consumption (214). Norwegian health authorities have enforced relatively strict public health initiatives regarding alcohol use, in accordance with available evidence on reducing alcohol problems (215). The availability of alcoholic beverages is low, alcohol taxes are high, and the drinking/driving legislation is strict and well enforced. The incomes of most physicians are above the population mean, so they are probably influenced by high alcohol prices to a lesser degree than other, economically challenged sections of the population. However, these public health interventions, which are known to affect drinkers at all consumption levels (216), may still go some way to explaining why the extent of hazardous drinking in our sample was below that in at least some other countries.

We found no decline in the prevalence of hazardous drinking after the physicians had left medical school. This finding is somewhat surprising, because studies of the general population have shown a decline in hazardous drinking in the early twenties (217;218). We would expect that leaving medical school, establishing a family, and entering professional

life would be paralleled by a decline in hazardous drinking. However, a study from the USA also found no reduction in drinking, and even an increase with age, among physicians and attorneys (5). Although our findings are partly consistent with the findings of Flaherty and Richman, they are difficult to explain and require further investigation. It is noteworthy that the lifestyles of physicians are generally reported to be healthier than those of the general population; for example, in terms of smoking, eating habits, and self-screening (13). Against this background, we would expect their alcohol consumption to be lower than that of the general population. However, the general impression derived from our results and the literature (30;32) is that the alcohol habits of physicians resemble those of the general population. This is perhaps reflected in the finding that physicians more often discuss smoking habits than alcohol consumption with their patients (219). Newbury-Birch (220), among others, maintains that we should expect better conduct from these individuals, who know all too well about the impact of alcohol on an individual's health. Furthermore, the issue of hazardous drinking is of particular relevance to medical students and physicians because they exert an influence disproportionate to their numbers on the social and economic health of their society.

Some studies have reported that female physicians in their middle and late careers use more alcohol than other women (5). Other data have also indicated that middle-aged (mean age, 41 years) female physicians consume alcohol more frequently than other working populations (221). We found no support for this among the female physicians in our sample, but this should be investigated further. Unpublished NORDOC data show a significantly higher occurrence of hazardous drinking among women in the Medical Student Cohort than in the six-year-older Young Doctor Cohort (11% vs 5%, respectively). This probably reflects the gradual shift in women's drinking habits towards those of men, which warrants concern and should also be studied further. Disturbingly, a recently published paper on American surgeons reported that the prevalence of alcohol abuse or dependence was almost twice as high for females (25.6%) than for males (13.9%) (222).

Although the extent of hazardous drinking among physicians may not exceed that of the general population, they may still constitute an especially vulnerable group. They report high levels of occupational stress (15;20) and stress related to the balancing of work and home obligations (72), and they may be at increased risk of affective disorders/depressive symptoms (4). Furthermore, physician rates of self-prescription are high (223), whereas their help-seeking rates are disturbingly low (224). When physicians develop alcohol problems, recognition and intervention are often delayed (225;226). Contrary to widespread

misconception, alcohol, rather than other substances, is the drug of choice among the majority (about 50%) of physicians who seek treatment for addiction (227;228). Alcohol problems can lead to further health problems among those who already suffer from depression (5;6). As mentioned in the Introduction, the most common dual diagnosis among physicians receiving psychological treatment is affective disorders and alcohol problems (229). In fact, it has been suggested that the rate of co-morbidity in physicians treated for substance use disorders has increased in recent years (229). The co-morbidity of alcohol abuse and depression amplifies an already elevated suicide risk among physicians (40). Finally, the potential negative consequences, not only for the individual physician but also for the health care that he or she provides, stress the importance of identifying vulnerability factors for hazardous drinking (1).

4.1.2 Can positive expectancy about the tension-reducing effects of alcohol predict hazardous drinking six years later?

We found that the expectancy that alcohol use reduces tension predicted hazardous drinking six years later, even when we controlled for hazardous drinking at baseline. This was the first study to investigate the long-term implications of alcohol expectancy on drinking behaviour in a representative sample. It was also the first study to examine alcohol expectancy among physicians. In keeping with our findings, Patrick and colleagues (230) reported in 2010 that alcohol expectancies in a national sample of adolescents predicted alcohol use as much as two decades later, even when the researchers controlled for earlier drinking behaviour. One might expect that physicians, who should be well aware of the negative consequences of hazardous drinking, would be less affected by such beliefs than other members of society. However, our results tell a different story. They provide strong evidence for the importance of such expectancies and underscore a potentially fruitful opportunity for interventions directed towards medical students and physicians.

It is especially interesting that we identified a predictive effect by applying only one of the six subscales that make up the AEQ. This is consistent with other studies that have reported that the perceived positive effects of alcohol are more influential than its perceived negative effects. Furthermore, the tension reduction subscale is the subscale that has been most consistently linked to problem drinking and alcohol dependence (231;232), which is also consistent with our results. Physicians report considerable occupational stress, and the tension reduction subscale is likely to be particularly important for individuals who experience excessive work stress. However, this requires further study.

The predictive effect of baseline hazardous drinking was high, even in the adjusted analysis. Male physicians had a 21 times higher risk of hazardous drinking at T2 if they reported this behaviour at T1. Some researchers have also suggested that drinking behaviour influences the development of expectancies, which in turn lead to further drinking. Our findings cannot rule this out, but a supplementary analysis indicated that the effect of hazardous drinking at T1 on hazardous drinking at T2 was only moderately reduced when alcohol expectancy was added into a logistic regression. Hence, if alcohol expectancy has a mediating effect on hazardous drinking, it is only small.

In summary, we found both hazardous drinking and the expectancy that alcohol use reduces tension, measured at the same time point, exerted independent effects on future hazardous drinking. Therefore, interventions that aim to curb hazardous drinking should focus on both expectancies (cognitions) and drinking behaviour.

4.1.3 Is the relationship between alcohol expectancy and hazardous drinking mediated by the use of alcohol to cope with stress?

We did not find support for the central tenet of social learning theory that the relationship between alcohol expectancy and hazardous drinking is mediated by the use of alcohol to cope with stress. Some other attempts to test this association have also failed to find a mediating role of the use of alcohol to cope (233).

Numerous studies have pointed to the many sources of stress on medical students (234-237) and physicians (15;20;238). Although a previous prospective study identified the use of alcohol to cope with stress as an independent risk factor for hazardous drinking among male Norwegian medical students (34), this finding was not repeated in the present study of Norwegian medical postgraduates. Others have suggested that medical students primarily report recreation or pleasure as motives for their drinking, rather than it being a way of coping with tension or anxiety (239). Perhaps the relatively low prevalence of alcohol use to cope (8% overall; 6% in women and 12% in men) reflects this. In any case, the one-item inventory may be less reliable than a multi-item inventory. This issue requires further study, perhaps with a broader measure of drinking to cope, to establish whether the model posed by social learning theory holds true. The association between stress and drinking behaviour among physicians, with the possible mediational or moderating effects of alcohol expectancy and use of alcohol to cope, also warrants further study. Finally, female-specific measures of hazardous drinking should be applied in all future studies of physicians' drinking behaviour to ensure that the results are also valid for women. It is

particularly pertinent to investigate female physicians' alcohol use over time because there is evidence that their alcohol use increases in mid-career stages.

4.2 Severe depressive symptoms among Norwegian physicians

4.2.1 Prevalence estimates of severe depressive symptoms throughout the 15-year follow-up period

The prevalence of severe depressive symptoms was as follows: 14% at T1, 8% at T2, 7% at T3, 10% at T4, and 7% at T5. There were no significant differences between the sexes in these estimates. There were significant declines from T1 to T2 and from T1 to T5. We have found no other studies in which the GHQ-28 *severe depression* subscale has been used alone, which makes any exact comparison difficult. However, a few studies have applied the complete GHQ-28 (all four subscales). Baldwin reported that 44% of medical students were identified as cases (38) when the 3/4 cut-off was used, which corresponds to our cut-off value of 1 or higher. Willcock also used the 3/4 cut-off on the complete GHQ-28 and reported that 26% of final-year medical students and 37% of interns were cases (240). Nash and colleagues found that 28% of Australian physicians (all age groups) were cases on the GHQ-28, but they applied a higher cut-off of 4/5 (212). Bruce (25) also applied a cut-off of 4/5 and found that 41% of older (mean age, 45 years) physicians were GHQ-28 cases. Because the scores on the specific subscales were not reported in any of the four studies mentioned here, we cannot compare them with our scores on the *severe depression* subscale, but the numbers reported seem somewhat higher than ours.

Our prevalence estimates also seem to be slightly lower than those reported in the literature from studies that have used instruments other than the GHQ-28 to measure depressive symptoms. Depressive symptoms were reported in 28% of British (20) and 23% of Canadian (21) interns ("case" prevalence among interns was 7.7% in the present study). Reuben found that 35% of first-year postgraduate residents on ward rotation experienced depressive symptoms (22). These differences can be explained by the fact that we examined *severe* depressive symptoms. The instruments used by Firth-Cozens (Hopkins Symptom Check List-10) and Hsu (Center for Epidemiologic Studies Depression Scale) may be more sensitive to distress, so the high prevalence estimates may reflect stress rather than mental-health problems. Furthermore, as mentioned earlier, the *severe depression* subscale may not identify milder depressive symptoms (241). Nevertheless, if Norwegian physicians experience less-intense depressive symptoms than other physician populations, how can we

explain this? We know that Norwegian physicians work fewer hours than physicians in some other countries, such as Germany (242). Residents in the USA work 80 hours or more a week (243), whereas Norwegian physicians work on average fewer than 50 hours (244). In a general population-based sample, long working hours were associated with the persistence of major depression (245), and sleep deprivation (80) predicted depression among physicians. However, the depressogenic effects of such work-related factors are not yet established (245;246) and remain to be investigated in longitudinal studies of Norwegian physicians.

Some studies have described a deterioration in medical students' mental health as they progress through medical school, particularly during the clinical curriculum (171;247), despite some inconsistencies in the literature (26;112), and the first year of residency seems to be associated with a further increase in emotional distress (76;248;249). Levels of depression have been reported to decline from the end of the first postgraduate year to the end of residency (76;248;249), but this does not seem to be in complete accord with the pattern identified here. Based on the literature, we suggest that a further increase rather than a decline occurs in the rate of depression from the final semester of medical school to internship. However, our observation point during internship was towards the end of the last semester, so the improvement in mood that has been described by others may have already started by the time we measured depression during internship. Moreover, the medical students responded to the survey in their graduating semester, just before their final exams, which may have induced a peak in emotional distress. The relatively stable rates from internship to the third, fourth, and fifth observation points are as expected, although as far as we know, there is no other such long-term longitudinal and representative study. It is also noteworthy that diagnostic interview studies of depressive symptoms have only been undertaken so far among medical students (71;250) and interns (79), together with a few studies of clinical samples (229). Until prevalence estimates are made based on diagnostic interviews with physicians beyond the very first years of their training, the incidence of depressive disorders in a representative sample of physicians remains unclear. Nevertheless, our results are encouraging. They show that physicians, as a group, experience less-severe depressive symptoms after medical school than during medical school. Studies of general population samples have reported that depressive symptoms tend to decrease from early adulthood to mid-life (251), which is consistent with the decline in the prevalence of severe depressive symptoms observed in the present study. An earlier analysis of NORDOC data found that work stress related to emotional pressure, time pressure, and fear of complaints

or criticism decreased during the 10 years following graduation (72). Although the same study found that the stress related to work–home interference increased in the same period, we can speculate that pressures related to work (emotionally demanding patients, time pressures, etc.) have stronger depressogenic effects. However, this remains to be studied. Overall, our longitudinal studies (NORDOC) have so far shown that Norwegian medical undergraduates and early postgraduates experience the highest levels of emotional distress on the physician career path. This suggests that any prevention or intervention should start early in the medical career, particularly because this distress may be associated with severe depression and suicidal ideation (91;107).

4.2.2 Can perceived parental bonding, measured at the end of medical school, predict severe depressive symptoms measured four and nine years after graduation from medical school?

We found that perceived low maternal care predicted severe depressive symptoms 4–9 years after graduating from medical school. Our study is the first to show the predictive value of parental bonding for depressive symptoms with such a long-term perspective in any population, and the first to study these relationships in a representative sample of physicians. The observations that 1) parental care is more important than overprotection (97;147), and 2) the perceived parental style of the mother is a more important predictor than that of the father also concur with the main impressions derived from earlier research (97;101;148;149;252;253). However, the literature is not completely consistent on this topic. In fact, a relatively recent longitudinal study by Overbeek and colleagues of a large (N = 5,000) general German population sample showed that lack of care and overprotection were equally important dimensions of parenting in the prediction of mental disorders (254). The authors also concluded that fathers and mothers were equally important in this respect. Despite somewhat incongruent findings, the results of the majority of studies to date are consistent with ours. Therefore, it seems wise to await replication of the German findings before we conclude otherwise. Mothers still spend more time with their children than fathers do (25), and although the pattern is gradually changing towards more paternal involvement, it seems likely that the effect of the mother's parenting is still stronger than that of the father's parenting. The respondents in Overbeek's study were not younger than the respondents in the present study, so the discrepancy is unlikely to reflect the gradually changing gender roles.

In conclusion, our findings about the relationship between parental bonding and depressive symptoms accord well with both earlier studies and psychological theory, as described on page 70, section 2. The two landmark prospective investigations of physicians, mentioned in the Introduction (page 19, section 3), both found that negative experiences in familial relationships predicted depression (28;78). However, the researchers did not control for personality in either study. The amount of variance in depressive symptoms that is explained by the perception of bonding is generally low at 1%–5% of the variance. Our complete model explained 16.5% of the variance, but when the perception of parental bonding was considered exclusively, the percentage was 7.7%. Hence, although the importance of perceived bonding experiences with caregivers seems indisputable, its extent should not be overstated. Interestingly, when these bonding perceptions are studied in interaction with, for example, life stress (negative life events), their combined influence may be considerably greater.

However, the independent effect of suboptimal parental bonding disappeared in an expanded multivariate model. This finding is discussed on page 73, section 2.

4.2.3 Is the relationship between parental bonding and depressive symptoms mediated by low general self-esteem?

We found partial support for the hypothesis that self-esteem mediates the observed relationship between perceived parental bonding and depressive symptoms. Eighty per cent of the effect of parental bonding on the mean severe depressive symptom score was mediated by low general self-esteem.

Our results are the first to demonstrate the mediating role of self-esteem between parental bonding and depressive symptoms with such a broad time window. Furthermore, our findings concur with the results of prospective studies of young physicians, which have demonstrated that both self-criticism (80) and neuroticism (79) predict depressive symptoms. Also consistent with our findings are previous demonstrations of an association between self-esteem and negative parenting, in both medical students (255) and other populations (100). Earlier attempts to test the mediation of self-esteem or neuroticism (256) have given inconsistent results (257-259). However, a relatively recent review of empirical research into the role of individuals' parenting as a developmental antecedent for depressive disorders concluded that negative cognitive styles may mediate this relationship (260).

A relevant question pertaining to our findings is whether an underlying element, such as a genetic factor, influences all the variables examined, producing reports of negative

bonding, low self-esteem, *and* severe depressive symptoms (261). This possibility notwithstanding, the parental bonding instrument has undergone thorough testing and has gained wide acceptance as a robust measure of perceived parenting. It is especially reassuring that the care dimension, in particular, which was the focus of the present study, has been found to withstand the influence of mood (145). This validates our findings, albeit based on self-report. We can also speculate whether there are higher levels of neuroticism in physicians than in other occupational groups, supporting the *vulnerability hypothesis*. One study indicated that this is so, because it found higher levels of neuroticism among Norwegian physicians than among Norwegian police officers, including when the sexes were compared separately (262).

Theoretically, our findings are consistent with the formulation of Blatt and colleagues, who described the influence of parent–child interactions on vulnerability to depression through the development of self-critical personality styles (99;263). One of the most influential depression models, Beck's cognitive theory of depression (98), is also consistent with our results. Negative self-schemas, or negatively loaded cognitive representations of the self, constitute the central vulnerability concept in this theory (264). In accordance with developmental psychological theories, Beck suggests that negative self-schemas are developed in caregiver relationships early in life. The child attempts to make sense of its experiences and concludes that negative interactions with caregivers result from his/her own deficiencies rather than from those of the caregiver. Our results are compatible with these theoretical suggestions. Interestingly, these considerations focus on a factor that may be important in, and amenable to, psychotherapy: students' and young physicians' tendency to self-criticize.

4.2.4 Can individual and stress-related factors, measured at the end of medical school, predict severe depressive symptoms in the 15 years after graduation from medical school?

Very few recent and representative studies have followed physicians into their careers after they leave medical school (238;265;266), and none for as long as 15 years. McManus and colleagues studied work-related factors and personality in relation to stress and burn-out (266). In their longitudinal study of Swiss residents, Buddeberg-Fischer and colleagues found work-related stress in the second year of residency to be a significant vulnerability factor for depression two years later (238). In a one-year prospective study of interns in the USA, Sen investigated other factors and found that the following predicted depressive

symptoms: difficult early environment, history of major depression, neuroticism, female sex, increased work hours, perceived medical errors, and stressful life events (198). They also found partial support for an interaction between stressful life events and the *5-HTTLPR* gene. West and his research group have also followed USA medical students into their early postgraduate years, but they have primarily investigated depression as a vulnerability factor for medical errors (265). Hence, the present study is the first to examine vulnerability factors for depressive symptoms in a representative longitudinal cohort over such a long follow-up period.

Generally, one cannot expect depressive symptoms among physicians to be caused by factors other than those seen among non-physicians. However, physicians constitute a selective subgroup of the population, and some studies have suggested that they may differ in some respects from the general population. (I) One study found that male medical students reported lower general self-esteem than the general population (3). (II) Two studies found that the occurrence of depressive disorders was higher among medical students than among the general population (71;250). (III) Other studies have reported high levels of compulsive behaviour and perfectionism among medical students (267;268). If these studies are valid, increased vulnerability to depressive symptoms might be expected in this occupational group. However, additional comparative studies of large representative samples are required to clarify this.

The predictors identified in our study were high reality weakness (OR = 2.3), high neuroticism (OR = 3.4), young age (OR = 1.1), and severe depressive symptoms at baseline (OR = 3.62). Our finding that reality weakness was an independent risk factor for subsequent severe depressive symptoms is original but to some extent consistent with the finding that personality psychopathology increases the risk of suicidal ideation (269). Reality weakness is assumed to reflect serious intrapsychic disturbance, and our findings underline the importance of identifying students who are troubled. Earlier publications based on the NORDOC data have found that this trait predicts suicidal ideation (91) and a lack of help-seeking, even when adjustment was made for the level of emotional distress (270).

Young age predicted severe depressive symptoms. We consider the most probable explanation for this is that individuals who begin medical school at an early age lack useful life experiences. They might be clever yet naïve and therefore less equipped to tackle the many challenges of medical school and subsequent working life than older students. In Norway today, most students enter medical school directly from high school, so they are

several years younger than the study sample. Consequently, they may be even more vulnerable to severe depressive symptoms than the subjects of our study. These findings suggest that the age of admission to medical school should be reconsidered. Would it be wise to delay entrance to the overwhelming medical curriculum by a couple of years?

Neuroticism has been established as a robust predictor of future depression in general population studies (271;272), and this finding concurs with previous (107) and recent studies (198) of interns. However, ours is the first study to show the predictive effect of this trait on depressive symptoms with such a broad time perspective. This underscores the importance of identifying neurotic students, who may be at increased risk of depression. However, physicians with a certain degree of neuroticism may make a particular contribution to health care. In fact, some research supports the notion that self-critical doctors have a greater capacity for empathy (20). Consequently, one may be safer in the hands of a self-critical and perfectionistic physician than in those of a more moderately tempered one, as long as neuroticism does not disable the physician and he/she is not overly self-conscious.

The independent effect of severe depressive symptoms at baseline is consistent with earlier research into depression. In fact, earlier depressive symptoms, together with female sex, have been found to be the most indisputable risk factors for depression (70). It is also in keeping with a previous study that demonstrated undergraduate depression to be a long-term predictor of depression in male physicians (80). However, to the best of our knowledge, the present study is the first to show this long-term effect among female physicians.

The following variables were not independent predictors of depressive symptoms in this study: perceived parental bonding, perceived medical school stress, perceived diagnostic and recording skills, and hazardous drinking. The failure of perceived parental bonding to predict severe depressive symptoms independently is perhaps not surprising. In paper II, we found that 80% of the effect of parental bonding was mediated by low self-esteem. Because the correlation between self-esteem and neuroticism is quite high, neuroticism probably absorbed the effect of parental bonding in the multivariate model. Therefore, this does not imply that parental bonding is unimportant in the development of depressive symptoms, but it emphasizes that personality is a strong and consistent predictor and may be one of several, unmeasured underlying variables that influence the other variables included in the present model. Although the effect of parental bonding is absorbed by other variables, our findings reported in paper II illustrate that it is a component in a

complex picture. Certainly, it may be important in the early development of an individual's susceptibility to depression.

Earlier studies have associated perceived medical school stress with anxiety and depressive symptoms (3;113), so it could be expected to predict severe depressive symptoms also. However, this type of stress may be predominantly associated with milder depressive symptoms, and therefore it was not significant in our multivariate analysis. In the same way, this may explain why the two self-perceived clinical skills, recording skills and diagnostic skills, failed to predict severe depressive symptoms. A recent paper from our department concluded that self-efficacy measures are severely limited in predicting the performance of actual communication skills (273), so they may be somewhat dubious variables because some individuals consistently overestimate, and others under-estimate, their skills. Consequently, these results are difficult to interpret.

Finally, we failed to identify a predictive effect of hazardous drinking on severe depressive symptoms. The literature is inconsistent regarding the direction of the alcohol–depression relationship. Some researchers have failed to identify a link between alcohol problems and depressive symptoms (274), whereas numerous studies have established that not only alcohol problems (115;275;276) but also binge drinking (277) are predictive of both depressive symptoms and suicidal behaviour (278). There is at least one possible explanation of our failure to find a predictive effect of hazardous drinking in the present study. Our decision to dichotomize the hazardous drinking variable at the median value (see page 37, section 2 in the *Methods section*) may imply that we have not accurately identified hazardous alcohol consumption. In support of this suggestion is the finding described under *Additional results* (see page 57, section 2) that only alcohol *dependence* was predictive of future severe depressive symptoms. Nevertheless, the association between hazardous drinking and severe depressive symptoms in the univariate analysis was almost significant ($p = 0.07$). Hence, rather than dismiss the possible depressogenic effect of hazardous drinking, further research should look into this with an appropriate measure, and particularly ensuring sex-specific sensitivity.

We also lack good comparative studies of depressive symptoms among physicians and academics. Future research should also encompass comprehensive predictor models of depressive symptoms. The relative importance of contextual (i.e., work-related stress and life stress) and individual factors has not been clarified. These issues should preferably be investigated with repeated-measures models to determine how their effects vary over time. It would also be interesting to study the links between genes and some of the identified risk

factors, such as reality weakness, which may have a role as possible endophenotypes with respect to mental-health problems (279). The gene × environment interaction could also be studied in NORDOC (198). Physicians are prone to relatively uniform stress levels in their jobs, and the interaction with genes and negative life events could be investigated in these two cohorts.

4.3 Sex-based issues

Men reported a higher prevalence of hazardous drinking than women at all observation points. This is consistent with other studies of medical students (33;36;280) and with the results of general population studies (281). However, we did not include a measure of hazardous drinking that was sensitive to female hazardous drinking, and the prevalence among women may be higher in reality than the estimates reported here. There is a tendency in the general population for the pattern of female drinking to approach that of males, and a comparison of the two NORDOC cohorts showed the same trend (34;87). Further longitudinal studies at more established stages of physicians' careers would shed more light upon this issue. As mentioned on page 15 (section 1), some studies have reported that female physicians in their middle and late careers use more alcohol than other women (30-32), and data from Norway Statistics indicate that middle-aged female physicians consume alcohol more frequently than other working populations (221). The pattern of female physicians' alcohol use should clearly be studied further in NORDOC.

No sex-based differences in the prevalence of severe depressive symptoms were identified at any of the five observation points. In fact, we have found negligible differences between the men and women in the NORDOC cohorts with respect to their affective symptoms or mental health (55;91;107;171). This is inconsistent with studies of the general population (281) and numerous studies of medical students (17;106) and physicians (38). The prevalence of depression is generally reported to be higher among the females. We found higher levels of medical school stress among female medical students at mid-curriculum but not at the time of graduation. Similarly, other studies of medical students and physicians have, like us, failed to identify any sex-based differences (26;80;239;282). There are three possible explanations for this lack of sex-based differences.

First, the mental health of male medical students and physicians is possibly weaker than that of men in the general population. In his early studies, Vaillant reported that male physicians were more likely than the controls to have relatively poor marriages, to use drugs

and alcohol more heavily, and to receive psychotherapy (28). Terman found that feelings of inferiority were more common among male physicians than other men (283). A more recent study found that male physicians had lower social functioning and vitality than comparable men in the general population (12), and Bramness and colleagues identified lower self-esteem among male medical students (3). Some data have also shown that more young male physicians consulted a psychologist or psychiatrist in the preceding 12 months than men in the general population (9).

Second, the lack of sex-based differences could also be caused by the mental health of women physicians exceeding that of their female counterparts in the general population. Our findings reported in paper II reveal that the association between perceived parental bonding and severe depressive symptoms is stronger among men. Consistent with these results, other research has suggested that high-achieving women maintain more cognitive distance from their parents and are more individualistic and internally motivated than other women (284). The female physicians in our study also reported a higher level of care by their fathers than male physicians. Hence, it can be inferred that "psychologically well-nourished" women self-select to the profession of medicine (282). On the contrary, *distress* levels are reportedly higher among female medical students than among male students (106;285), and the predictive effect of neuroticism on perceived job stress in the first postgraduate year was stronger for the women in our sample (166). However, our measure of severe depressive symptoms is not sensitive to distress, and the female physicians may still be over-represented in terms of mild depressive symptoms. Interestingly, a recent publication reporting NORDOC data from the 10 first postgraduate years identified no sex-based differences in stress regarding the work–home interface (72). That study also suggested that women more often reduce their working hours and gather more social support than men (286). Perhaps this shows that female physicians protect themselves against the possible detrimental effects of stress more than male physicians, which may partly explain why the prevalence of depressive symptoms is not elevated in female physicians. Consistent with this, the elevated self-prescription rates among young male physicians relative to those of their female counterparts may also reflect suboptimal coping among these men.

Third, this lack of differences may be caused by both the better mental health of the women physicians and the poorer health among the men relative to their counterparts in the general population, thus narrowing the sex-based difference in mental health. This may be attributable to less-rigid sex-based roles among medical students and young physicians, with

the women being more masculine than other women and the men being more care oriented than other men. This androgynous shift may explain the lack of differences between the sexes, but it remains to be studied. Our results support this third explanation, that the lack of sex-based differences in the prevalence of severe depressive symptoms among physicians is the result of a combination of the two processes discussed above.

4.4 Summary of the main findings

In this study, we have found support for the hypothesis suggested by social learning theory that positive expectancy about the tension-reducing effects of alcohol can predict hazardous drinking. Another tenet of social learning theory, that this effect is mediated by the actual use of alcohol to cope, was rejected. Consistent with both cognitive and psychoanalytical theory, suboptimal parenting was found to predict severe depressive symptoms. This relationship was partly explained by low self-esteem. In an expanded multivariate model, reality weakness and neuroticism, in addition to young age, were independent risk factors for severe depressive symptoms.

4.5 Implications for medical education and the promotion of mental health

Our finding that the prevalence of depressive symptoms is higher in medical school than during the postgraduate years suggests that early intervention is important. An earlier NORDOC study found an elevated prevalence of suicidal ideation among medical students (91). In another NORDOC study, 30% of medical students reported mental-health problems requiring treatment, for which only a minority had sought help (171). The issues of stigmatization and fear of later “professional sanctions” are assumed to be associated with this low help-seeking behaviour (132). Against the background of our findings, we discuss possible changes that could be implemented.

4.5.1 Admission to medical school

Young age was predictive of later depressive symptoms. Therefore, we should ask whether young students are prepared for the many pressures experienced in medical school. Our findings imply that the admission age to medical school should be reconsidered. Would it be wise to delay entrance to the overwhelming medical curriculum by a couple of years? In any case, it seems wise to educate students about what to expect and how to cope with the many stresses of medical education. Whether or not personality traits should be used as selection

criteria might also be an issue worth considering. Previously, a NORDOC study concluded that interpersonal or therapeutic support is possibly more important than selection based on "normal" personality traits (109). Nevertheless, applicants who score highly for reality weakness may benefit from career advice that prompts them towards an option other than the medical profession, but this requires further investigation.

4.5.2 The medical curriculum

Depressive symptoms and hazardous drinking among medical students and physicians could also be introduced as a topic in the medical curriculum. Some researchers doubt that medical students are sufficiently aware of the negative impact of alcohol. The medical curriculum at the University of Oslo includes nearly two days of lectures that focus directly on alcohol problems (287). It is tempting to speculate that this important topic could usefully be given greater emphasis. Similarly, several researchers have made similar suggestions about the medical education in their respective countries (288;289). Others have expressed concern that medical students have minimal exposure to the topic of physician health at all (239;290;291). Following this line of argument, medical students should be encouraged to seek professional help if they encounter problems. If their role models display non-judgemental, positive attitudes towards help-seeking, it is likely that the stigma associated with mental-health problems will gradually decrease. Students should be introduced to these topics and given the opportunity to reflect on and discuss them. Some researchers have suggested that medical students should be introduced to recovering physicians (292;293), a suggestion that we support.

Because the stigma related to mental-health problems still exists, many physicians avoid seeking help because they fear a loss of confidentiality. It may be difficult to seek help without recognition by someone in the health system, especially in small Norwegian communities. In a small country like Norway, we must strive to preserve physicians' confidentiality.

Finally, teachers involved in medical training should be aware of the factors that increase their students' vulnerability to depressive symptoms and hazardous drinking. A low level of perceived maternal care predicts severe depressive symptoms, and this association is partly mediated by low self-esteem. Relating to other people is a central component of the physician's work. Therefore, we can speculate that young physicians with low self-esteem, originating from negative experiences with caregivers, may be especially vulnerable to the

challenging and sometimes humiliating environments of medical school and early clinical practice.

4.5.3 Postgraduate training

Our findings have revealed that the prevalence of hazardous drinking does not decline among physicians during the postgraduate years and that positive expectancy of the tension-reducing properties of alcohol is a risk factor at this level. Postgraduate training should focus on both alcohol expectancy and hazardous drinking. A recent intervention study found that a single-session group-delivered programme significantly reduced alcohol expectancy and subsequent hazardous alcohol use among university students (294). Such results are promising, and this strategy could also be considered for resident doctors.

Although we identified a reduction in the prevalence of severe depressive symptoms during their postgraduate years, we found that low self-esteem predicted these symptoms during this period. The senior physicians who encounter physicians in residency should openly acknowledge that this is a stressful time in the young physicians' lives, which can challenge their self-esteem. The residents should also be encouraged to seek help if they are struggling. By explicitly addressing this topic, the threshold for help-seeking is likely to be lowered.

4.5.4 Interventions

Some students are likely to benefit from help in managing their feelings of imperfection and lack of approval, and in reducing their high levels of self-criticism. Two personality traits, neuroticism and reality weakness, were more potent predictors of severe depressive symptoms on the univariate analysis than perceived medical school stress, the two clinical competence variables (diagnostic and recording skills), or parental bonding. Although personality traits are known to have a considerable degree of heritability and are therefore presumably stable, they are amenable to some change (295;296).

A literature review concluded that stress management interventions such as mindfulness training (285) are very effective in medical education (297). Self-development groups are sufficient for some students (298), whereas others may need individual psychotherapy that focuses on self-esteem issues (299).

Although research into the outcomes of physicians is not methodologically optimal, reports generally conclude that the recovery rates of physicians are better than those of the

general population. More than 90% of the physicians who receive psychological treatment for alcohol problems are working again five years later (300). Studies have also shown that the medical school and early residency stages of training are ideal times for intervention (301).

4.5.5 Health services for physicians in Norway

A special health service for physicians has already been established in Norway. The idea behind the so-called "Physician for physician" arrangement (*Lege-for-lege-ordningen*), which was established in the early 1990s, is that physicians can contact physicians specifically suited to attending physician patients (224). However, it seems that this arrangement is well run in only four of Norway's 19 provinces (personal communication, Tone H. Holter at Legeforeningen). Furthermore, the amount and type of training that the physicians' physicians receive are somewhat arbitrary. We suggest that the training of these specialist physicians be standardized, preferably with a focus on the mental-health problems of physicians.

Villa Sana (Modum Bad, Norway) is a resource centre in Norway that offers a counselling programme designed to improve the mental health and quality of life of physicians. The self-referred counselling intervention, which aims to motivate physicians to reflect on their own situations and personal needs, has produced a reduction in burn-out (302). We are enthusiastic about the therapeutic work conducted at Villa Sana and recommend that it be made known to medical students and physicians at all stages of their careers.

REFERENCE LIST

(1) Strang J, Wilks M, Wells B, Marshall J. Missed problems and missed opportunities for addicted doctors. BMJ 1998;316(7129):405–6.

(2) Toyry S, Rasanen K, Kujala S, Aarimaa M, Juntunen J, Kalimo R, et al. Self-reported health, illness, and self-care among Finnish physicians: a national survey. Arch Fam Med 2000;9(10):1079–85.

(3) Bramness JG, Fixdal TC, Vaglum P. Effect of medical school stress on the mental health of medical students in early and late clinical curriculum. Acta Psychiatr Scand 1991;84(4):340–5.

(4) Firth-Cozens J. Depression in doctors. In: Robertson MM, Katona CLE, editors. Depression and physical illness. New York: John Wiley & Sons Ltd; 1997. p. 95–111.

(5) Flaherty JA, Richman JA. Substance use and addiction among medical students, residents, and physicians. Psychiatr Clin North Am 1993;16(1):189–97.

(6) Center C, Davis M, Detre T, Ford DE, Hansbrough W, Hendin H, et al. Confronting depression and suicide in physicians: a consensus statement. JAMA 2003;289(23):3161–6.

(7) Isaksson R. Emotional exhaustion and distress after a counselling intervention for physicians: a three-year prospective longitudinal cohort study. PhD thesis, University of Oslo, Oslo, 2009.

(8) Rovik JO. The role of personality in stress, burnout and help-seeking. A ten-year longitudinal study among Norwegian medical students and early career physicians. PhD thesis, University of Oslo, Oslo, 2009.

(9) Tyssen R. Mental health problems among medical students and young physicians. A nationwide and longitudinal study. PhD thesis, University of Oslo, Oslo, 2001.

(10) Hem E. Suicidal behaviour in some human service occupations with special emphasis on physicians and police: a nationwide study. PhD thesis, University of Oslo, Oslo, 2004.

(11) McKevitt C, Morgan M, Dundas R, Holland WW. Sickness absence and ‘working through’ illness: a comparison of two professional groups. J Public Health Med 1997;19(3):295–300.

(12) Stavem K, Hofoss D, Aasland OG, Loge JH. The self-perceived health status of Norwegian physicians compared with a reference population and foreign physicians. Scand J Public Health 2001;29(3):194–9.

(13) Frank E, Brogan DJ, Mokdad AH, Simoes EJ, Kahn HS, Greenberg RS. Health-related behaviors of women physicians vs other women in the United States. Arch Intern Med 1998;158(4):342–8.

(14) Wall TD, Bolden RI, Borrill CS, Carter AJ, Golya DA, Hardy GE, et al. Minor psychiatric disorder in NHS trust staff: occupational and gender differences. Br J Psychiatry 1997;171:519–23.

(15) Falkum E, Gjerberg E, Hofoss D, Aasland OG. [Time stress among Norwegian physicians]. Tidsskr Nor Laegeforen 1997;117(7):954–9.

(16) Tyssen R, Hem E, Gude T, Gronvold NT, Ekeberg O, Vaglum P. Lower life satisfaction in physicians compared with a general population sample: a 10-year longitudinal, nationwide study of course and predictors. Soc Psychiatry Psychiatr Epidemiol 2009;44(1):47–54.

(17) Dyrbye LN, Thomas MR, Shanafelt TD. Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students. Acad Med 2006;81(4):354–73.

(18) Clark DC, Zeldow PB. Vicissitudes of depressed mood during four years of medical school. JAMA 1988;260(17):2521–8.

(19) Schwenk TL, Davis L, Wimsatt LA. Depression, stigma, and suicidal ideation in medical students. JAMA 2010;304(11):1181–90.

(20) Firth-Cozens J. Emotional distress in junior house officers. Br Med J (Clin Res Ed) 1987;295(6597):533–6.

(21) Hsu K, Marshall V. Prevalence of depression and distress in a large sample of Canadian residents, interns, and fellows. Am J Psychiatry 1987;144(12):1561–6.

(22) Reuben DB. Depressive symptoms in medical house officers. Effects of level of training and work rotation. Arch Intern Med 1985;145(2):286–8.

(23) Firth-Cozens J. Predicting stress in general practitioners: 10 year follow up postal survey. BMJ 1997;315(7099):34–5.

(24) Caplan RP. Stress, anxiety, and depression in hospital consultants, general practitioners, and senior health service managers. BMJ 1994;309(6964):1261–3.

(25) Bruce CT, Sanger MM, Thomas PS, Petkus JR, Yates DH. Factors affecting female or male consultant stress in an Australian teaching hospital. Med J Aust 2003;179(3):174–5.

(26) Katz ED, Sharp L, Ferguson E. Depression among emergency medicine residents over an academic year. Acad Emerg Med 2006;13(3):284–7.

(27) Ford DE, Mead LA, Chang PP, Cooper-Patrick L, Wang NY, Klag MJ. Depression is a risk factor for coronary artery disease in men: the Precursors Study. Arch Intern Med 1998;158(13):1422–6.

(28) Vaillant GE, Sobowale NC, McArthur C. Some psychologic vulnerabilities of physicians. N Engl J Med 1972;287(8):372–5.

(29) Murray RM. Alcoholism amongst male doctors in Scotland. Lancet 1976;2(7988):729–31.

(30) Hughes PH, Brandenburg N, Baldwin DC Jr, Storr CL, Williams KM, Anthony JC, et al. Prevalence of substance use among US physicians. JAMA 1992;267(17):2333–9.

(31) Brewster JM. Prevalence of alcohol and other drug problems among physicians. JAMA 1986;255(14):1913–20.

(32) Gulbrandsen P, Aasland OG. [Changes in drinking habits among Norwegian physicians 1985–2000]. Tidsskr Nor Laegeforen 2002;122(29):2791–4.

(33) Newbury-Birch D, Walshaw D, Kamali F. Drink and drugs: from medical students to doctors. Drug Alcohol Depend 2001;64(3):265–70.

(34) Kjobli J, Tyssen R, Vaglum P, Aasland O, Gronvold NT, Ekeberg O. Personality traits and drinking to cope as predictors of hazardous drinking among medical students. J Stud Alcohol 2004;65(5):582–5.

(35) Collier DJ, Beales IL. Drinking among medical students: a questionnaire survey. BMJ 1989;299(6690):19–22.

(36) Dahlin M, Nilsson C, Stotzer E, Runeson B. Mental distress, alcohol use and help-seeking among medical and business students: a cross-sectional comparative study. BMC Med Educ 2011;11(1):92.

(37) Aasland OG, Amundsen A, Bruusgaard D, Jervell J, Morland J. [Alcohol drinking habits among Norwegian physicians]. Tidsskr Nor Laegeforen 1987;107(29):2553–8.

(38) Baldwin PJ, Dodd M, Wrate RM. Young doctors' health—II. Health and health behaviour. Soc Sci Med 1997;45(1):41–4.

(39) Frank E, Biola H, Burnett CA. Mortality rates and causes among U.S. physicians. Am J Prev Med 2000;19(3):155–9.

(40) Schernhammer ES, Colditz GA. Suicide rates among physicians: a quantitative and gender assessment (meta-analysis). Am J Psychiatry 2004;161(12):2295–302.

(41) Hem E, Haldorsen T, Aasland OG, Tyssen R, Vaglum P, Ekeberg O. Suicide among physicians. Am J Psychiatry 2005;162(11):2199–200.

(42) Lindeman S, Laara E, Vuori E, Lonnqvist J. Suicides among physicians, engineers and teachers: the prevalence of reported depression, admissions to hospital and contributory causes of death. Acta Psychiatr Scand 1997;96(1):68–71.

(43) Kessler RC, Berglund P, Borges G, Nock M, Wang PS. Trends in suicide ideation, plans, gestures, and attempts in the United States, 1990–1992 to 2001–2003. JAMA 2005;293(20):2487–95.

(44) Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. Arch Gen Psychiatry 1999;56(7):617–26.

(45) Henriksson MM, Aro HM, Marttunen MJ, Heikkinen ME, Isometsa ET, Kuoppasalmi KI, et al. Mental disorders and comorbidity in suicide. Am J Psychiatry 1993;150(6):935–40.

(46) Isometsa ET, Henriksson MM, Aro HM, Heikkinen ME, Kuoppasalmi KI, Lonnqvist JK. Suicide in major depression. Am J Psychiatry 1994;151(4):530–6.

(47) Roy A, Linnoila M. Alcoholism and suicide. Suicide Life Threat Behav 1986;16(2):244–73.

(48) Day, NL. Epidemiology of alcohol use, abuse, and dependence. In: Tsuang M, Tohen M, Zahner G, editors. Textbook in psychiatric epidemiology. New York: John Wiley & Sons; 1995. p. 345–60.

(49) Cheng AT, Chen TH, Chen CC, Jenkins R. Psychosocial and psychiatric risk factors for suicide. Case–control psychological autopsy study. Br J Psychiatry 2000;177:360–5.

(50) Roness A, Kaldestad E. [Mental disorders among physicians hospitalized in a psychiatric clinic]. Tidsskr Nor Laegeforen 1991;111(30):3619–22.

(51) Clayton PJ, Marten S, Davis MA, Wochnik E. Mood disorder in women professionals. J Affect Disord 1980;2(1):37–46.

(52) Roy A. Suicide in doctors. Psychiatr Clin North Am 1985;8(2):377–87.

(53) Tillett R. The patient within—psychopathology in the helping professions. Adv Psychiatr Treat 2003;9:272–9.

(54) Rosvold EO, Bjertness E. Illness behaviour among Norwegian physicians. Scand J Public Health 2002;30(2):125–32.

(55) Tyssen R, Vaglum P, Gronvold NT, Ekeberg O. The impact of job stress and working conditions on mental health problems among junior house officers. A nationwide Norwegian prospective cohort study. Med Educ 2000;34(5):374–84.

(56) Tyssen R, Vaglum P, Gronvold NT, Ekeberg O. Factors in medical school that predict postgraduate mental health problems in need of treatment. A nationwide and longitudinal study. Med Educ 2001;35(2):110–20.

(57) Brooke D, Edwards G, Taylor C. Addiction as an occupational hazard: 144 doctors with drug and alcohol problems. Br J Addict 1991;86(8):1011–6.

(58) Institute of Medicine. Broadening the base of treatment for alcohol problems. Washington DC; National Academy Press; 1990.

(59) Aasland O. Trends in smoking and drinking among Norwegian physicians. In Proceedings of the International Conference on Physician Health. 1994. Ottawa, Canada.

(60) Frank E, Rothenberg R, Lewis C, Belodoff BF. Correlates of physicians' prevention-related practices. Findings from the Women Physicians' Health Study. Arch Fam Med 2000;9(4):359–67.

(61) Young AS, Klap R, Sherbourne CD, Wells KB. The quality of care for depressive and anxiety disorders in the United States. Arch Gen Psychiatry 2001;58(1):55–61.

(62) Luoma JB, Martin CE, Pearson JL. Contact with mental health and primary care providers before suicide: a review of the evidence. Am J Psychiatry 2002;159(6):909–16.

(63) Ford S, Lewis S, Fallowfield L. Psychological morbidity in newly referred patients with cancer. J Psychosom Res 1995;39(2):193–202.

(64) Mann JJ, Apter A, Bertolote J, Beautrais A, Currier D, Haas A, et al. Suicide prevention strategies: a systematic review. JAMA 2005;294(16):2064–74.

(65) Lockley SW, Cronin JW, Evans EE, Cade BE, Lee CJ, Landrigan CP, et al. Effect of reducing interns' weekly work hours on sleep and attentional failures. N Engl J Med 2004;351(18):1829–37.

(66) West CP, Tan AD, Habermann TM, Sloan JA, Shanafelt TD. Association of resident fatigue and distress with perceived medical errors. JAMA 2009;302(12):1294–300.

(67) Fahrenkopf AM, Sectish TC, Barger LK, Sharek PJ, Lewin D, Chiang VW, et al. Rates of medication errors among depressed and burnt out residents: prospective cohort study. BMJ 2008;336(7642):488–91.

(68) O'Hagan J, Richards J. In sickness and in health. 2nd ed. Wellington (New Zealand) Doctors Health Advisory Service (DHAS); 1998;1-9.

(69) Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. Lancet 2009;374(9702):1714–21.

(70) Ingram RE, Price JM. Vulnerability to psychopathology: risk across the lifespan. 2nd ed. New York, London: The Guilford Press; 2010.

(71) Zoccolillo M, Murphy GE, Wetzel RD. Depression among medical students. J Affect Disord 1986;11(1):91–6.

(72) Rovik JO, Tyssen R, Hem E, Gude T, Ekeberg O, Moum T, et al. Job stress in young physicians with an emphasis on the work–home interface: a nine-year, nationwide and longitudinal study of its course and predictors. Ind Health 2007;45(5):662–71.

(73) Kjeldstadli K, Tyssen R, Finset A, Hem E, Gude T, Gronvold NT, et al. Life satisfaction and resilience in medical school—a six-year longitudinal, nationwide and comparative study. BMC Med Educ 2006;6:48.

(74) Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am J Psychiatry 2008;165(10):1272–80.

(75) Munoz RF, Cuijpers P, Smit F, Barrera AZ, Leykin Y. Prevention of major depression. Annu Rev Clin Psychol 2010;6:181–212.

(76) Girard DE, Hickam DH, Gordon GH, Robison RO. A prospective study of internal medicine residents' emotions and attitudes throughout their training. Acad Med 1991;66(2):111–4.

(77) Firth-Cozens J. The role of early family experiences in the perception of organizational stress: fusing clinical and organizational perspectives. J Occupat Organizat Psychol 1992;65(1):61–75.

(78) Thomas CB, Duszynski KR. Closeness to parents and the family constellation in a prospective study of five disease states: suicide, mental illness, malignant tumor, hypertension and coronary heart disease. Johns Hopkins Med J 1974;134(5):251–70.

(79) Clark DC, Salazar-Grueso E, Grabler P, Fawcett J. Predictors of depression during the first 6 months of internship. Am J Psychiatry 1984;141(9):1095–8.

(80) Firth-Cozens J. Individual and organizational predictors of depression in general practitioners. Br J Gen Pract 1998;48(435):1647–51.

(81) Institute of Medicine. Broadening the base of treatment for alcohol problems: report of a study by a committee of the Institute of Medicine, Division of Mental Health and Behavioral Medicine. Washington DC: National Academy Press; 1990.

(82) Tuunanen M, Aalto M, Seppa K. Binge drinking and its detection among middle-aged men using AUDIT, AUDIT-C and AUDIT-3. Drug Alcohol Rev 2007;26(3):295–9.

(83) Saitz R. Clinical practice. Unhealthy alcohol use. N Engl J Med 2005;352(6):596–607.

(84) Fleming MF, Barry KL, MacDonald R. The Alcohol Use Disorders Identification Test (AUDIT) in a college sample. Int J Addict 1991;26(11):1173–85.

(85) Rotter J. Level of aspiration as a method of studying personality. II. Development and evaluation of a controlled method. J Exper Psychol 1942;31:410–22.

(86) Cooper ML, Russell M, George WH. Coping, expectancies, and alcohol abuse: a test of social learning formulations. J Abnorm Psychol 1988;97(2):218–30.

(87) Tyssen R, Vaglum P, Aasland OG, Gronvold NT, Ekeberg O. Use of alcohol to cope with tension, and its relation to gender, years in medical school and hazardous drinking: a study of two nation-wide Norwegian samples of medical students. Addiction 1998;93(9):1341–9.

(88) Moore RD, Mead L, Pearson TA. Youthful precursors of alcohol abuse in physicians. Am J Med 1990;88(4):332–6.

(89) Stacy AW, Newcomb MD, Bentler PM. Cognitive motivation and drug use: a 9-year longitudinal study. J Abnorm Psychol 1991;100(4):502–15.

(90) Goldberg D, Williams P. A user's guide to the General Health Questionnaire. Windsor, UK: NFER-Nelson; 1988.

(91) Tyssen R, Vaglum P, Gronvold NT, Ekeberg O. Suicidal ideation among medical students and young physicians: a nationwide and prospective study of prevalence and predictors. J Affect Disord 2001;64(1):69–79.

(92) Freudenberger H. Burnout: contemporary issues, trends, and concerns. In: Farber B, editor. Stress and burnout in the human service professions. New York: Pergamon Press; 1983. p. 23–8.

(93) Miller A. Prisoners of childhood: the drama of the gifted child and the search for the true self. New York: Basic Books; 1979.

(94) Johnson WD. Predisposition to emotional distress and psychiatric illness amongst doctors: the role of unconscious and experiential factors. Br J Med Psychol 1991;64(4):317–29.

(95) Whiffen VE, Parker GB, Wilhelm K, Mitchell PB, Malhi G. Parental care and personality in melancholic and nonmelancholic depression. J Nerv Ment Dis 2003;191(6):358–64.

(96) Oakley-Browne MA, Joyce PR, Wells JE, Bushnell JA, Hornblow AR. Adverse parenting and other childhood experience as risk factors for depression in women aged 18–44 years. J Affect Disord 1995;34(1):13–23.

(97) Enns MW, Cox BJ, Larsen DK. Perceptions of parental bonding and symptom severity in adults with depression: mediation by personality dimensions. Can J Psychiatry 2000;45(3):263–8.

(98) Beck AT, Rush AJ, Shaw BF, Emery G. Cognitive therapy of depression. New York: The Guilford Press; 1979.

(99) Blatt SJ, Homann E. Parent–child interaction in the etiology of dependent and self-critical depression. Clin Psychol Rev 1992;12(1):47–91.

(100) Parker G. Parental rearing style: examining for links with personality vulnerability factors for depression. Soc Psychiatry Psychiatr Epidemiol 1993;28(3):97–100.

(101) Rey JM. Perceptions of poor maternal care are associated with adolescent depression. J Affect Disord 1995;34(2):95–100.

(102) Gladstone GL, Parker GB. The role of parenting in the development of psychopathology: an overview of research using the Parental Bonding Instrument. In: Hudson JL, Rapee RM, editors. Psychopathology and the family. New York (NY): Elsevier; 2005. p. 21–33.

(103) Baldwin PJ, Dodd M, Wrate RW. Young doctors' health—I. How do working conditions affect attitudes, health and performance? Soc Sci Med 1997;45(1):35–40.

(104) Tyssen R, Vaglum P. Mental health problems among young doctors: an updated review of prospective studies. Harv Rev Psychiatry 2002;10(3):154–65.

(105) Goldney RD, Dal GE, Fisher LJ, Wilson D. Population attributable risk of major depression for suicidal ideation in a random and representative community sample. J Affect Disord 2003;74(3):267–72.

(106) Dahlin M, Joneborg N, Runeson B. Stress and depression among medical students: a cross-sectional study. Med Educ 2005;39(6):594–604.

(107) Tyssen R, Hem E, Vaglum P, Gronvold NT, Ekeberg O. The process of suicidal planning among medical doctors: predictors in a longitudinal Norwegian sample. J Affect Disord 2004;80(2–3):191–8.

(108) Vaglum P, Falkum E. Self-criticism, dependency and depressive symptoms in a nationwide sample of Norwegian physicians. J Affect Disord 1999;52(1–3):153–9.

(109) Tyssen R, Dolatowski FC, Rovik JO, Thorkildsen RF, Ekeberg O, Hem E, et al. Personality traits and types predict medical school stress: a six-year longitudinal and nationwide study. Med Educ 2007;41(8):781–7.

(110) Kapur N, Borrill C, Stride C. Psychological morbidity and job satisfaction in hospital consultants and junior house officers: multicentre, cross sectional survey. BMJ 1998;317(7157):511–2.

(111) Ramirez AJ, Graham J, Richards MA, Cull A, Gregory WM. Mental health of hospital consultants: the effects of stress and satisfaction at work. Lancet 1996;347(9003):724–8.

(112) Hainer BL, Palesch Y. Symptoms of depression in residents: a South Carolina Family Practice Research Consortium study. Acad Med 1998;73(12):1305–10.

(113) Vitaliano PP, Russo J, Carr JE, Heerwagen JH. Medical school pressures and their relationship to anxiety. J Nerv Ment Dis 1984;172(12):730–6.

(114) Chandavarkar U, Azzam A, Mathews CA. Anxiety symptoms and perceived performance in medical students. Depress Anxiety 2007;24(2):103–11.

(115) Fergusson DM, Boden JM, Horwood LJ. Tests of causal links between alcohol abuse or dependence and major depression. Arch Gen Psychiatry 2009;66(3):260–6.

(116) Saunders JB, Aasland OG, Babor TF, de la Fuente Jr, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption—II. Addiction 1993;88(6):791–804.

(117) Reinert DF, Allen JP. The Alcohol Use Disorders Identification Test (AUDIT): a review of recent research. Alcohol Clin Exp Res 2002;26(2):272–9.

(118) Allen JP, Litten RZ, Fertig JB, Babor T. A review of research on the Alcohol Use Disorders Identification Test (AUDIT). Alcohol Clin Exp Res 1997;21(4):613–9.

(119) Fiellin DA, Reid MC, O'Connor PG. Screening for alcohol problems in primary care: a systematic review. Arch Intern Med 2000;160(13):1977–89.

(120) Allen JP, Reinert DF, Volk RJ. The alcohol use disorders identification test: an aid to recognition of alcohol problems in primary care patients. Prev Med 2001;33(5):428–33.

(121) Bradley KA, Kivlahan DR, Bush KR, McDonell MB, Fihn SD. Variations on the CAGE alcohol screening questionnaire: strengths and limitations in VA general medical patients. Alcohol Clin Exp Res 2001;25(10):1472–8.

(122) Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998;158(16):1789–95.

(123) Bradley KA, Bush KR, Epler AJ, Dobie DJ, Davis TM, Sporleder JL, et al. Two brief alcohol-screening tests from the Alcohol Use Disorders Identification Test (AUDIT): validation in a female Veterans Affairs patient population. Arch Intern Med 2003;163(7):821–9.

(124) Wechsler H, Davenport A, Dowdall G, Moeykens B, Castillo S. Health and behavioral consequences of binge drinking in college. A national survey of students at 140 campuses. JAMA 1994;272(21):1672–7.

(125) Goldberg DP. The detection of psychiatric illness by questionnaire. London: Oxford University Press; 1972.

(126) Lobo A, Perez-Echeverria MJ, Artal J. Validity of the scaled version of the General Health Questionnaire (GHQ-28) in a Spanish population. Psychol Med 1986;16(1):135–40.

(127) Goldberg DP, Hillier VF. A scaled version of the General Health Questionnaire. Psychol Med 1979;9(1):139–45.

(128) Koeter MW. Validity of the GHQ and SCL anxiety and depression scales: a comparative study. J Affect Disord 1992;24(4):271–9.

(129) Goldman R, Robinson D, Grube B, Hanks R, Putnam K, Walder, et al. General psychiatric symptom measures. In: Rush AJ, Pincus H, First M, Blacker D, Endicott J, Keith S, et al., editors. Handbook of psychiatric measures. 1st ed. Washington DC: American Psychiatric Association; 2000. p. 71–92.

(130) Goldberg DP, Oldehinkel T, Ormel J. Why GHQ threshold varies from one place to another. Psychol Med 1998;28(4):915–21.

(131) Willmott SA, Boardman JA, Henshaw CA, Jones PW. Understanding General Health Questionnaire (GHQ-28) score and its threshold. Soc Psychiatry Psychiatr Epidemiol 2004;39(8):613–7.

(132) Chew-Graham CA, Rogers A, Yassin N. 'I wouldn't want it on my CV or their records': medical students' experiences of help-seeking for mental health problems. Med Educ 2003;37(10):873–80.

(133) Brown SA, Christiansen BA, Goldman MS. The Alcohol Expectancy Questionnaire: an instrument for the assessment of adolescent and adult alcohol expectancies. J Stud Alcohol 1987;48(5):483–91.

(134) Leigh BC. Confirmatory factor analysis of alcohol expectancy scales. J Stud Alcohol 1989;50(3):268–77.

(135) George WH, Frone MR, Cooper ML, Russell M, Skinner JB, Windle M. A revised Alcohol Expectancy Questionnaire: factor structure confirmation, and invariance in a general population sample. J Stud Alcohol 1995;56(2):177–85.

(136) West R. Theory of Addiction. London: Blackwell Publishing; Addiction Press; 2006.

(137) Timmer SG, Verhoff J, Colten ME. Life stress, helplessness, and the use of alcohol and drugs to cope: an analysis of national survey data. In: Shiffman S, Wills TA, editors. Coping and substance use. Orlando, FL: Academic Press; 1985. p. 171–98.

(138) Sterud T, Hem E, Ekeberg O, Lau B. Occupational stress and alcohol use: a study of two nationwide samples of operational police and ambulance personnel in Norway. J Stud Alcohol Drugs 2007;68(6):896–904.

(139) Dybek I, Bischof G, Grothues J, Reinhardt S, Meyer C, Hapke U, et al. The reliability and validity of the Alcohol Use Disorders Identification Test (AUDIT) in a German general practice population sample. J Stud Alcohol 2006;67(3):473–81.

(140) Bergman H, Kallmen H, Rydberg U, Sandahl C. [Ten questions about alcohol as identifier of addiction problems. Psychometric tests at an emergency psychiatric department]. Lakartidningen 1998;95(43):4731–5.

(141) Aertgeerts B, Buntinx F, Ansoms S, Fevery J. Screening properties of questionnaires and laboratory tests for the detection of alcohol abuse or dependence in a general practice population. Br J Gen Pract 2001;51(464):206–17.

(142) Hammersley R. A digest of memory phenomena for addiction research. Addiction 1994;89(3):283–93.

(143) Horverak Ø, Bye E. Det norske drikkemønsteret; en studie basert på intervjudata fra 1973–2004. Oslo: Statens Institutt for Rusmiddelforskning; 2007. Report No. 2.

(144) Parker G, Tupling H, Brown LB. A parental bonding instrument. Br J Med Psychol 1979;52(1):1–10.

(145) Wilhelm K, Boyce P, Brownhill S. The relationship between interpersonal sensitivity, anxiety disorders and major depression. J Affect Disord 2004;79(1–3):33–41.

(146) Lewinsohn PM, Rosenbaum M. Recall of parental behavior by acute depressives, remitted depressives, and nondepressives. J Pers Soc Psychol 1987;52(3):611–9.

(147) Duggan C, Sham P, Minne C, Lee A, Murray R. Quality of parenting and vulnerability to depression: results from a family study. Psychol Med 1998;28(1):185–91.

(148) Mackinnon A, Henderson AS, Andrews G. Parental 'affectionless control' as an antecedent to adult depression: a risk factor refined. Psychol Med 1993;23(1):135–41.

(149) Parker G, Hadzi-Pavlovic D, Greenwald S, Weissman M. Low parental care as a risk factor to lifetime depression in a community sample. J Affect Disord 1995;33(3):173–80.

(150) Rodgers B. Reported parental behaviour and adult affective symptoms. 1. Associations and moderating factors. Psychol Med 1996;26(1):51–61.

(151) Gotlib IH, Mount JH, Cordy NI, Whiffen VE. Depression and perceptions of early parenting: a longitudinal investigation. Br J Psychiatry 1988;152:24–7.

(152) Wilhelm K, Parker G. Reliability of the parental bonding instrument and intimate bond measure scales. Aust N Z J Psychiatry 1990;24(2):199–202.

(153) Murphy E, Wickramaratne P, Weissman M. The stability of parental bonding reports: a 20-year follow-up. J Affect Disord 2010.125(1-3):307-15.

(154) Nitta M, Narita T, Umeda K, Hattori M, Naitoh H, Iwata N. Influence of negative cognition on the parental bonding instrument (PBI) in patients with major depression. J Nerv Ment Dis 2008;196(3):244–6.

(155) Parker G. The Parental Bonding Instrument: psychometric properties reviewed. Psychiatr Dev 1989;7(4):317–35.

(156) Parker G. Validating an experimental measure of parental style: the use of a twin sample. Acta Psychiatr Scand 1989;73:22–7.

(157) Torgersen S. Hereditary–environmental differentiation of general neurotic, obsessive, and impulsive hysterical personality traits. Acta Genet Med Gemellol (Roma) 1980;29(3):193–207.

(158) Torgersen S, Alnæs R. Localizing DSM-III personality disorders in a three-dimensional structural space. J Personal Disord 1989;3:274–81.

(159) Lazare A, Klerman GL, Armor DJ. Oral, obsessive, and hysterical personality patterns. An investigation of psychoanalytic concepts by means of factor analysis. Arch Gen Psychiatry 1966;14(6):624–30.

(160) Neilands TB, Silvera DH, Perry JA, Richardsen A, Holte A. A validation and short form of the Basic Character Inventory. Scand J Psychol 2008;49(2):161–8.

(161) Torgersen S. The oral, obsessive, and hysterical personality syndromes. A study of hereditary and environmental factors by means of the twin method. Arch Gen Psychiatry 1980;37(11):1272–7.

(162) Alnaes R, Torgersen S. Basic character inventory personality traits among patients with major depression, anxiety disorders and mixed conditions. Eur Arch Psychiatry Neurol Sci 1990;239(5):303–8.

(163) Ravndal E, Vaglum P. Changes in antisocial aggressiveness during treatment in a hierarchical therapeutic community. A prospective study of personality changes. Acta Psychiatr Scand 1991;84(6):524–30.

(164) Perry J, Silvera D, Rosenvinge JH, Holte A. Are oral, obsessive, and hysterical personality traits related to disturbed eating patterns? A general population study of 6,313 men and women. J Pers Assess 2002;78(3):405–16.

(165) Aasland OG, Olff M, Falkum E, Schweder T, Ursin H. Health complaints and job stress in Norwegian physicians: the use of an overlapping questionnaire design. Soc Sci Med 1997;45(11):1615–29.

(166) Tyssen R, Vaglum P, Gronvold NT, Ekeberg O. The relative importance of individual and organizational factors for the prevention of job stress during internship: a nationwide and prospective study. Med Teach 2005;27(8):726–31.

(167) Eysenck HJ. The biological basis of personality. Springfield, IL: Charles C. Thomas; 1967.

(168) Cloninger CR. A systematic method for clinical description and classification of personality variants. A proposal. Arch Gen Psychiatry 1987;44(6):573–88.

(169) Costa P, McCrae R. Personality disorders and the five-factor model of personality. J Pers Dis. 1990;4:362–371.

(170) Costa PJ, McCrae R. NEO PI-R Professional Manual. Odessa, FL: Psychological Assessment Resources, Inc.; 1992.

(171) Midtgaard M, Ekeberg O, Vaglum P, Tyssen R. Mental health treatment needs for medical students: a national longitudinal study. Eur Psychiatry 2008;23(7):505–11.

(172) Ward M, Gruppen L, Regehr G. Measuring self-assessment: current state of the art. Adv Health Sci Educ Theory Pract 2002;7(1):63–80.

(173) Blanch-Hartigan D. Medical students' self-assessment of performance: results from three meta-analyses. Patient Educ Couns 2011;84(1):3–9.

(174) Bakken LL, Sheridan J, Carnes M. Gender differences among physician-scientists in self-assessed abilities to perform clinical research. Acad Med 2003;78(12):1281–6.

(175) Leopold SS, Morgan HD, Kadel NJ, Gardner GC, Schaad DC, Wolf FM. Impact of educational intervention on confidence and competence in the performance of a simple surgical task. J Bone Joint Surg Am 2005;87(5):1031–7.

(176) Depue RA, Monroe SM. Conceptualization and measurement of human disorder in life stress research: the problem of chronic disturbance. Psychol Bull 1986;99(1):36–51.

(177) Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986;51(6):1173–82.

(178) Diggle PJ, Liang KY, Zeger ZL. Analysis of longitudinal data. Oxford, England: Clarendon Press; 1994.

(179) Little R, Rubin D. Statistical analysis with missing data. New York: Wiley; 2002.

(180) Mazumdar S, Tang G, Houck PR, Dew MA, Begley AE, Scott J, et al. Statistical analysis of longitudinal psychiatric data with dropouts. J Psychiatr Res 2007;41(12):1032–41.

(181) de Lange AH, Taris TW, Kompier MA, Houtman IL, Bongers PM. The relationships between work characteristics and mental health: examining normal, reversed and reciprocal relationships in a 4-wave study. Work Stress 2004;18:149–166.

(182) Dormann C, Zapf D. Social stressors at work, irritation, and depressive symptoms: accounting for unmeasured third variables in a multi-wave study. J Occupat Organizat Psychol 2002;75:33–58.

(183) Moum T. Is subjective well-being a predictor of nonresponse in broad population surveys? Soc Indic Res 1994;32:1–20.

(184) Berk ML. Interviewing physicians: the effect of improved response rate. Am J Public Health 1985;75(11):1338–40.

(185) Knudsen AK, Hotopf M, Skogen JC, Overland S, Mykletun A. The health status of nonparticipants in a population-based health study: the Hordaland Health Study. Am J Epidemiol 2010;172(11):1306–14.

(186) Wainer H. Inferences from self-selected samples. New York: Springer Verlag; 1986.

(187) Gibbins R, Israel Y, Kalant H, Popham RE, Schmidt W, Smart R. Research advances in alcohol and drug problems, Vol. 1. New York: Wiley; 1974.

(188) Burgess AM Jr, Tierney JT. Bias due to nonresponse in a mail survey of Rhode Island physicians' smoking habits—1968. N Engl J Med 1970;282(16):908.

(189) Lautenschlager GFV. Computer administration of questions: more desirable or more social desirability? J Appl Psychol 1990;75:310–4.

(190) Paulhus DL. Two-component models of socially desirable responding. J Pers Soc Psychol 1984;46:598–609.

(191) Moum T. Mode of administration and interviewer effects in self-reported symptoms of anxiety and depression. Soc Indic Res 1998;45:279–318.

(192) Midanik L. The validity of self-reported alcohol consumption and alcohol problems: a literature review. Br J Addict 1982;77(4):357–82.

(193) Midanik L. Over-reports of recent alcohol consumption in a clinical population: a validity study. Drug Alcohol Depend 1982;9(2):101–10.

(194) Midanik LT. Validity of self-reported alcohol use: a literature review and assessment. Br J Addict 1988;83(9):1019–30.

(195) Rouse B, Kozel N, Richards L. Self report methods of estimating drug abuse: meeting current challenges to validity. Research Monograph. Rockville, MD, National Clearinghouse for Alcohol and Drug Information. 1985.

(196) Hilton ME. A comparison of a prospective diary and two summary recall techniques for recording alcohol consumption. Br J Addict 1989;84(9):1085–92.

(197) Midanik LT. Perspectives on the validity of self-reported alcohol use. Br J Addict 1989;84(12):1419–23.

(198) Sen S, Kranzler HR, Krystal JH, Speller H, Chan G, Gelernter J, et al. A prospective cohort study investigating factors associated with depression during medical internship. Arch Gen Psychiatry 2010;67(6):557–65.

(199) Kellerman SE, Herold J. Physician response to surveys. A review of the literature. Am J Prev Med 2001;20(1):61–7.

(200) Martin B. Don't survey physicians! Chicago, IL: Center for Health Services Research and Development, American Medical Association; 1974.

(201) Asch DA, Jedrziewski MK, Christakis NA. Response rates to mail surveys published in medical journals. J Clin Epidemiol 1997;50(10):1129–36.

(202) Cartwright A. Professionals as responders: variations in and effects of response rates to questionnaires, 1961–77. Br Med J 1978;2(6149):1419–21.

(203) Cartwright A, Windsor J. Identification of people attending outpatients: congruity between hospital records and response to a postal screen. Community Med 1989;11(3):225–9.

(204) Compton MT, Frank E. Mental health concerns among Canadian physicians: results from the 2007–2008 Canadian Physician Health Study. Compr Psychiatry 2011;52(5):542–7.

(205) Shanafelt TD, Balch CM, Dyrbye L, Bechamps G, Russell T, Satele D, et al. Special report: suicidal ideation among American surgeons. Arch Surg 2011;146(1):54–62.

(206) McCarty CA, Kosterman R, Mason WA, McCauley E, Hawkins JD, Herrenkohl TI, et al. Longitudinal associations among depression, obesity and alcohol use disorders in young adulthood. Gen Hosp Psychiatry 2009;31(5):442–50.

(207) Clark DC. Alcohol and drug use and mood disorders among medical students: implications for physician impairment. QRB Qual Rev Bull 1988;14(2):50–4.

(208) Khantzian EJ. The self-medication hypothesis of addictive disorders: focus on heroin and cocaine dependence. Am J Psychiatry 1985;142(11):1259–64.

(209) Rist F, Glockner-Rist A, Demmel R. The Alcohol Use Disorders Identification Test revisited: establishing its structure using nonlinear factor analysis and identifying subgroups of respondents using latent class factor analysis. Drug Alcohol Depend 2009;100(1–2):71–82.

(210) Juntunen J, Asp S, Olkinuora M, Aarimaa M, Strid L, Kauttu K. Doctors' drinking habits and consumption of alcohol. BMJ 1988;297(6654):951–4.

(211) Rosta J. Hazardous alcohol use among hospital doctors in Germany. Alcohol Alcohol 2008;43(2):198–203.

(212) Nash LM, Daly MG, Kelly PJ, van Ekert EH, Walter G, Walton M, et al. Factors associated with psychiatric morbidity and hazardous alcohol use in Australian doctors. Med J Aust 2010;193(3):161–6.

(213) Kenna GA, Wood MD. Alcohol use by healthcare professionals. Drug Alcohol Depend 2004;75(1):107–16.

(214) Horverak Ø, Nordlund S, Rossow I. Om sentrale deler av norsk alkoholpolitikk. Oslo: Statens Institutt for Rusmiddelforskning (SIRUS); 2001.

(215) Babor T, Caetano R, Casswell S, Edwards G. Alcohol: no ordinary commodity: research and public policy. Oxford and London: Oxford University Press; 2010.

(216) Dawson D. Defining risk drinking. Alcohol Res Health 2011;34(2):144–56.

(217) Donovan JE, Jessor R, Jessor L. Problem drinking in adolescence and young adulthood. A follow-up study. J Stud Alcohol 1983;44(1):109–37.

(218) Hammer T, Vaglum P. Use of alcohol and drugs in the transitional phase from adolescence to young adulthood. J Adolesc 1990;13(2):129–42.

(219) Nygaard P, Paschall MJ, Aasland OG, Lund KE. Use and barriers to use of screening and brief interventions for alcohol problems among Norwegian general practitioners. Alcohol Alcohol 2010;45(2):207–12.

(220) Newbury-Birch D, Walshaw D, Kamali F. Drink and drugs: from medical students to doctors. Drug Alcohol Depend 2001;64(3):265–70.

(221) Aasland O. Unpublished data from Legeforeningens forskningsinstitutt. 2005.

(222) Oreskovich MR, Kaups KL, Balch CM, Hanks JB, Satele D, Sloan J, et al. Prevalence of alcohol use disorders among American surgeons. Arch Surg 2012;147(2):168–74.

(223) Hem E, Stokke G, Tyssen R, Gronvold NT, Vaglum P, Ekeberg O. Self-prescribing among young Norwegian doctors: a nine-year follow-up study of a nationwide sample. BMC Med 2005;3:16.

(224) Tyssen R. Health problems and the use of health services among physicians: a review article with particular emphasis on Norwegian studies. Ind Health 2007;45(5):599–610.

(225) Talbott GD, Gallegos KV, Wilson PO, Porter TL. The Medical Association of Georgia's Impaired Physicians Program. Review of the first 1000 physicians: analysis of specialty. JAMA 1987;257(21):2927–30.

(226) Brooke D, Edwards G, Andrews T. Doctors and substance misuse: types of doctors, types of problems. Addiction 1993;88(5):655–63.

(227) Berge KH, Seppala MD, Schipper AM. Chemical dependency and the physician. Mayo Clin Proc 2009;84(7):625–31.

(228) Buhl A, Oreskovich MR, Meredith C, Campbell M, DuPont RL. Prognosis for the recovery of surgeons from chemical dependency: a 5-year outcome study. Arch Surg 2011;146:1286–91.

(229) Lusilla P, Gual A, Roncero C, Bruguera E, Marcos V, Valero S, et al. Dual diagnosis in inpatient physicians: prevalence and clinical characteristics. Ment Health Subst Use 2008;1:10–20.

(230) Patrick ME, Wray-Lake L, Finlay AK, Maggs JL. The long arm of expectancies: adolescent alcohol expectancies predict adult alcohol use. Alcohol Alcohol 2010;45(1):17–24.

(231) Brown SA. Expectancies versus background in the prediction of college drinking patterns. J Consult Clin Psychol 1985;53(1):123–30.

(232) Mooney DK, Fromme K, Kivlahan DR, Marlatt GA. Correlates of alcohol consumption: sex, age, and expectancies relate differentially to quantity and frequency. Addict Behav 1987;12(3):235–40.

(233) Read JP, Wood MD, Kahler CW, Maddock JE, Palfai TP. Examining the role of drinking motives in college student alcohol use and problems. Psychol Addict Behav 2003;17(1):13–23.

(234) Ford CV, Wentz DK. Internship: what is stressful? South Med J 1986;79(5):595–9.

(235) Girard DE, Elliot DL, Hickam DH, Sparr L, Clarke NG, Warren L, et al. The internship—a prospective investigation of emotions and attitudes. West J Med 1986;144(1):93–8.

(236) Richman JA, Flaherty JA, Rospenda KM, Christensen ML. Mental health consequences and correlates of reported medical student abuse. JAMA 1992;267(5):692–4.

(237) Frank E, Carrera JS, Stratton T, Bickel J, Nora LM. Experiences of belittlement and harassment and their correlates among medical students in the United States: longitudinal survey. BMJ 2006;333(7570):682.

(238) Buddeberg-Fischer B, Klaghofer R, Stamm M, Siegrist J, Buddeberg C. Work stress and reduced health in young physicians: prospective evidence from Swiss residents. Int Arch Occup Environ Health 2008;82(1):31–8.

(239) Newbury-Birch D, White M, Kamali F. Factors influencing alcohol and illicit drug use amongst medical students. Drug Alcohol Depend 2000;59(2):125–30.

(240) Willcock SM, Daly MG, Tennant CC, Allard BJ. Burnout and psychiatric morbidity in new medical graduates. Med J Aust 2004;181(7):357–60.

(241) Malt UF, Mogstad TE, Refnin IB. [Goldberg's General Health Questionnaire]. Tidsskr Nor Laegeforen 1989;109(13):1391–4.

(242) Rosta J, Aasland OG. Work hours and self rated health of hospital doctors in Norway and Germany. A comparative study on national samples. BMC Health Serv Res 2011;11:40.

(243) Oransky I. Limits imposed on residents' work hours. Reduced workload said to increase patient—and resident—safety. Lancet 2003;362(9381):378–9.

(244) Stavem K, Hofoss D, Aasland OG. Work characteristics and morbidity as predictors of self-perceived health status in Norwegian physicians. Scand J Public Health 2003;31(5):375–81.

(245) Wang JL, Patten SB, Currie S, Sareen J, Schmitz N. Predictors of 1-year outcomes of major depressive disorder among individuals with a lifetime diagnosis: a population-based study. Psychol Med 2012;42(2):327-34.

(246) Netterstrom B, Conrad N, Bech P, Fink P, Olsen O, Rugulies R, et al. The relation between work-related psychosocial factors and the development of depression. Epidemiol Rev 2008;30:118–32.

(247) Hojat M, Mangione S, Nasca TJ, Rattner S, Erdmann JB, Gonnella JS, et al. An empirical study of decline in empathy in medical school. Med Educ 2004;38(9):934–41.

(248) Buddeberg-Fischer B, Klaghofer R, Buddeberg C. [Stress at work and well-being in junior residents]. Z Psychosom Med Psychother 2005;51(2):163–78.

(249) Bellini LM, Shea JA. Mood change and empathy decline persist during three years of internal medicine training. Acad Med 2005;80(2):164–7.

(250) Dahlin ME, Runeson B. Burnout and psychiatric morbidity among medical students entering clinical training: a three year prospective questionnaire and interview-based study. BMC Med Educ 2007;7:6.

(251) Jorm AF. Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult life span. Psychol Med 2000;30(1):11–22.

(252) Enns MW, Cox BJ, Clara I. Parental bonding and adult psychopathology: results from the US National Comorbidity Survey. Psychol Med 2002;32(6):997–1008.

(253) Hall LA, Peden AR, Rayens MK, Beebe LH. Parental bonding: a key factor for mental health of college women. Issues Ment Health Nurs 2004;25(3):277–91.

(254) Overbeek G, ten Have M, Vollebergh W, de Graf R. Parental lack of care and overprotection. Longitudinal associations with DSM-III-R disorders. Soc Psychiatry Psychiatr Epidemiol 2007;42(2):87–93.

(255) Brewin CR, Firth-Cozens J, Furnham A, McManus C. Self-criticism in adulthood and recalled childhood experience. J Abnorm Psychol 1992;101(3):561–6.

(256) Sato T, Uehara T, Narita T, Sakado K, Fujii Y. Parental bonding and personality in relation to a lifetime history of depression. Psychiatry Clin Neurosci 2000;54(2):121–30.

(257) McCranie EW, Bass JD. Childhood family antecedents of dependency and self-criticism: implications for depression. J Abnorm Psychol 1984;93(1):3–8.

(258) Matheson K, Kelly O, Cole B, Tannenbaum B, Dodd C, Anisman H. Parental bonding and depressive affect: the mediating role of coping resources. Br J Soc Psychol 2005;44(Pt 3):371–95.

(259) Lloyd C, Miller PM. The relationship of parental style to depression and self-esteem in adulthood. J Nerv Ment Dis 1997;185(11):655–63.

(260) Alloy LB, Abramson LY, Smith JM, Gibb BE, Neeren AM. Role of parenting and maltreatment histories in unipolar and bipolar mood disorders: mediation by cognitive vulnerability to depression. Clin Child Fam Psychol Rev 2006;9(1):23–64.

(261) McGuffin P, Perroud N, Uher R, Butler A, Aitchison KJ, Craig I, et al. The genetics of affective disorder and suicide. Eur Psychiatry 2010;25(5):275–7.

(262) Berg AM, Hem E, Lau B, Ekeberg O. An exploration of job stress and health in the Norwegian police service: a cross sectional study. J Occup Med Toxicol 2006;1:26.

(263) Blatt SJ, Wein SJ, Chevron ES, Quinlan DM. Parental representations and depression in normal young adults. J Abnorm Psychol 1979;88(4):388–97.

(264) Beck AT. Depression: clinical, experimental, and theoretical aspects. New York: Harper & Row; 1967.

(265) West CP, Huschka MM, Novotny PJ, Sloan JA, Kolars JC, Habermann TM, et al. Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study. JAMA 2006;296(9):1071–8.

(266) McManus IC, Keeling A, Paice E. Stress, burnout and doctors' attitudes to work are determined by personality and learning style: a twelve year longitudinal study of UK medical graduates. BMC Med 2004;2:29.

(267) Gabbard GO. The role of compulsiveness in the normal physician. JAMA 1985;254(20):2926–9.

(268) Krakowski AJ. Stress and the practice of medicine. III. Physicians compared with lawyers. Psychother Psychosom 1984;42(1–4):143–51.

(269) Brezo J, Paris J, Tremblay R, Vitaro F, Zoccolillo M, Hebert M, et al. Personality traits as correlates of suicide attempts and suicidal ideation in young adults. Psychol Med 2006;36(2):191–202.

(270) Tyssen R, Rovik JO, Vaglum P, Gronvold NT, Ekeberg O. Help-seeking for mental health problems among young physicians: is it the most ill that seeks help?—A longitudinal and nationwide study. Soc Psychiatry Psychiatr Epidemiol 2004;39(12):989–93.

(271) Kendler KS, Gatz M, Gardner CO, Pedersen NL. Personality and major depression: a Swedish longitudinal, population-based twin study. Arch Gen Psychiatry 2006;63(10):1113–20.

(272) Mattisson C, Bogren M, Horstmann V, Tambs K, Munk-Jorgensen P, Nettelbladt P. Risk factors for depressive disorders in the Lundby cohort—a 50 year prospective clinical follow-up. J Affect Disord 2009;113(3):203–15.

(273) Finset A, Anvik T, Bærheim A, Fasmer O, Grimstad H, Nordøy T, et al. Can young physicians' self efficacy regarding communication skills predict their performance in consultations with simulated patients before and after internship? Submitted for publication.

(274)Moscato BS, Russell M, Zielezny M, Bromet E, Egri G, Mudar P, et al. Gender differences in the relation between depressive symptoms and alcohol problems: a longitudinal perspective. Am J Epidemiol 1997;146(11):966–74.

(275) Hasin DS, Grant BF. Major depression in 6050 former drinkers: association with past alcohol dependence. Arch Gen Psychiatry 2002;59(9):794–800.

(276) Wang J, Patten SB. Prospective study of frequent heavy alcohol use and the risk of major depression in the Canadian general population. Depress Anxiety 2002;15(1):42–5.

(277) Paljarvi T, Koskenvuo M, Poikolainen K, Kauhanen J, Sillanmaki L, Makela P. Binge drinking and depressive symptoms: a 5-year population-based cohort study. Addiction 2009;104(7):1168–78.

(278) Borges G, Loera CR. Alcohol and drug use in suicidal behaviour. Curr Opin Psychiatry 2010;23(3):195-204.

(279) Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 2003;160(4):636–45.

(280) Richman JA, Flaherty JA. Sex differences in drinking among medical students: patterns and psychosocial correlates. J Stud Alcohol 1986;47(4):283–9.

(281) Kringlen E, Torgersen S, Cramer V. A Norwegian psychiatric epidemiological study. Am J Psychiatry 2001;158(7):1091–8.

(282) Richman JA, Flaherty JA. Gender differences in medical student distress: contributions of prior socialization and current role-related stress. Soc Sci Med 1990;30(7):777–87.

(283) Terman L. Scientists and non-scientists in a group of 800 gifted men. Psychol Monogr. 1954;68:1–44.

(284) Tangri SS. Determinants of occupational role innovation among college women. J Social Issues 1972;28(2):177–99.

(285) De Vibe M, Solhaug I, Tyssen R, Bjørndal A, Rosenvinge J, Sørlie T, et al. Mindfulness training for stress management: a randomised controlled study of Norwegian medical and psychology students. Submitted for publication.

(286) Gross CP, Mead LA, Ford DE, Klag MJ. Physician, heal thyself? Regular source of care and use of preventive health services among physicians. Arch Intern Med 2000 Nov 27;160(21):3209–14.

(287) Waal H. Personal communication. 2011.

(288) McAuliffe WE, Rohman M, Santangelo S, Feldman B, Magnuson E, Sobol A, et al. Psychoactive drug use among practicing physicians and medical students. N Engl J Med 1986;315(13):805–10.

(289) Krishel S, Richards CF. Alcohol and substance abuse training for emergency medicine residents: a survey of U.S. programs. Acad Emerg Med 1999;6(9):964–6.

(290) Croen LG, Woesner M, Herman M, Reichgott M. A longitudinal study of substance use and abuse in a single class of medical students. Acad Med 1997;72(5):376–81.

(291) Weissberg M. The meagerness of physicians' training in emergency psychiatric intervention. Acad Med 1990;65(12):747–50.

(292) Lewis DC, Faggett WL. Policy report of the Physician Consortium on Substance Abuse Education. 1991. Publ. no. HRSA-P-DM-91-3, U.S. Department of Health and Human Services, Bureau of Health Professions.

(293) Quilty LC, Meusel LA, Bagby RM. Neuroticism as a mediator of treatment response to SSRIs in major depressive disorder. J Affect Disord 2008 Nov;111(1):67–73.

(294) Fried AB, Dunn ME. The Expectancy Challenge Alcohol Literacy Curriculum (ECALC): a single session group intervention to reduce alcohol use. Psychol Addict Behav 2012. In press.

(295) Tang TZ, DeRubeis RJ, Hollon SD, Amsterdam J, Shelton R, Schalet B. Personality change during depression treatment: a placebo-controlled trial. Arch Gen Psychiatry 2009;66(12):1322–30.

(296) Lambert MJ. Bergin and Garfiled's Handbook of Psychotherapy and Behavior Change. New York: John Wiley and Sons, Inc.; 2004.

(297) Shapiro SL, Shapiro DE, Schwartz GE. Stress management in medical education: a review of the literature. Acad Med 2000;75(7):748–59.

(298) Holm M, Tyssen R, Stordal KI, Haver B. Self-development groups reduce medical school stress: a controlled intervention study. BMC Med Educ 2010;10:23.

(299) Fennell MJV. Depression, low self-esteem and mindfulness. Behav Res Ther 2004;42(9):1053–67.

(300) Mansky PA. Issues in the recovery of physicians from addictive illnesses. Psychiatr Q 1999;70(2):107–22.

(301) Hughes PH, Baldwin DC Jr, Sheehan DV, Conard S, Storr CL. Resident physician substance use, by specialty. Am J Psychiatry 1992;149(10):1348–54.

(302) Ro KE, Gude T, Tyssen R, Aasland OG. Counselling for burnout in Norwegian doctors: one year cohort study. BMJ 2008;337:a2004.

(303) Mykletun A, Bjerkeset O, Overland S, Prince M, Dewey M, Stewart R. Levels of anxiety and depression as predictors of mortality: the HUNT study. Br J Psychiatry 2009;195(2):118–25.

(304) Skogen JC, Knudsen AK, Mykletun A, Nesvag S, Overland S. Alcohol consumption, problem drinking, abstention and disability pension award. The Nord-Trondelag Health Study (HUNT). Addiction 2011;107(1):98-108.

5 PAPERS

Paper I

Kjersti s. Grotmol, Per Vaglum, Øivind Ekeberg, Tore Gude Olaf G. Aasland and Reidar Tyssen

Alcohol expectancy and hazardous drinking: a 6-year longitudinal and nationwide study of medical doctors

European Addiction Research 2010 16:17–22.

Paper II

Kjersti S. Grotmol, Øivind Ekeberg, Arnstein Finset, Tore Gude, Torbjørn Moum, Per Vaglum and Reidar Tyssen

Parental bonding and self-esteem as predictors of severe depressive symptoms: a 10-year follow-up study of Norwegian physicians

The Journal of Nervous and Mental Disease 2010 198:22–27.

Paper III

Kjersti S. Grotmol, Tore Gude, Torbjørn Moum, Per Vaglum and Reidar Tyssen

Risk factors at medical school for later severe depression: a 15-year longitudinal, nationwide study (NORDOC)*

Submitted for publication

*The Longitudinal Study of Norwegian Medical Students and Doctors

I

II

III

Risk factors at medical school for later severe depression: A 15-year longitudinal, nationwide study (NORDOC)*

Authors: Kjersti Støen Grotmol (PhD student)[1;2], Tore Gude (PhD)[1], Torbjørn Moum (PhD)[1], Per Vaglum (PhD)[1] & Reidar Tyssen (PhD)[1]

[1]Department of Behavioural Sciences in Medicine, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, PO Box 1111 Blindern, 0317 Oslo, Norway.

[2]Corresponding author: Kjersti Støen Grotmol. Telephone: + 47 22 85 10 20. E-mail address: k.s.grotmol@medisin.uio.no.

Word count: 3,622

* The Longitudinal Study of Norwegian Medical Students and Doctors

ABSTRACT

Background: Doctors have an increased risk of suicide, and depressive symptoms are prominent among young doctors. We lack prospective studies that identify risk factors that may be targeted in medical schools.

Methods: From 1993 to 2008, graduating medical students (n = 631) from all four Norwegian universities participated in the Longitudinal Study of Norwegian Medical Students and Doctors (NORDOC). After their graduating term (T1), they were followed up 1 (T2), 4 (T3), 10 (T4), and 15 (T5) years later with postal surveys. Severe depressive symptoms were measured by the General Health Questionnaire-28 (T1, T2, T3, T4, and T5) and analyzed by Generalized Estimating Equations.

Results: At T1 and T5, 13.7% and 7.2%, respectively, of the doctors reported severe depressive symptoms; a significant reduction over time (p = 0.001) in both genders. Independent risk factors for future depressive symptoms were: high neuroticism (odds ratio [OR] 3.4, 95% confidence interval [CI] 1.5–7.6, p = 0.003); high reality weakness (OR 2.3, 95% CI 1.2–4.2, p = 0.008); low age (OR 1.1, 95% CI 1.04–1.2, p = 0.003); and severe depressive symptoms at T1 (OR 3.6, 95% CI 2.1–6.1, $p < 0.001$).

Limitations: Selection bias and concurrent life and work stress may have influenced the results.

Conclusions: In addition to low age, high neuroticism yielded a 3 fold increased risk over the 15-year follow-up, whereas high reality weakness, which is linked to personality pathology, doubled the risk. These factors are clinically relevant for identification of students at risk.

1. Introduction

Doctors are at increased risk of suicide compared with other academics and health professionals (Schernhammer and Colditz, 2004; Hem et al., 2005), and depressive symptoms are prominent among medical students (Dyrbye et al., 2006) and young doctors (Sen et al., 2010). Therefore, studies that establish the risk factors for depression in doctors are required to identify individuals already at risk in medical school so that they can be offered support and treatment (Center et al., 2003). Individual characteristics such as personality traits, in particular neuroticism, may be important (Tyssen and Vaglum, 2002). Our research group has previously identified a personality characteristic related to personality pathology ('reality weakness') as a risk factor for suicidal ideation (Tyssen et al., 2004a). Other possible risk factors for depression are female sex (Sen et al., 2010) and a difficult early environment, such as poor parental bonding (Grotmol et al., 2010). Medical school-specific factors, such as perceived medical school stress (Vitaliano et al., 1984) and perceived clinical competence and skills (Chandavarkar et al., 2007), may influence the risk of depression, as may alcohol problems (Firth-Cozens, 1998). All such factors should be included as predictors in a longitudinal study. In the present study, we aimed to identify risk factors that can be measured in medical school. Therefore, contextual factors such as work stress and negative life events were not included in the present study, although they are known to influence the risk of depression. This is the first nationwide, long-term prospective study of doctors to identify possible early risk factors for severe depressive symptoms beyond the very first postgraduate years. We used data from a 15-year longitudinal study of Norwegian graduating medical students (NORDOC; $n = 631$), assessed five times for depressive symptoms, to obtain the odds ratios for individual and/or medical school-related factors, measured in the final term of medical school, that predict future severe depressive symptoms.

2. Methods

2.1. Sample and study design

The Young Doctor Cohort of the Longitudinal Study of Norwegian Medical Students and Doctors (NORDOC) consisted of all students who graduated in 1993 and 1994 from all four medical faculties in Norway (n = 631) (Tyssen et al., 2000). The respondents were surveyed by postal questionnaire at five time points: T1, the final year of medical school (n = 522/631, 83% of the eligible sample); T2, internship, 1 year after graduation (n = 371/631, 56% of the eligible sample); T3, the fourth postgraduate year (n = 422/631, 67% of the eligible sample); T4, the tenth postgraduate year (n = 390/631, 62% of the eligible sample); and T5, the 15th postgraduate year (n = 330/631, 52% of the eligible sample). Two hundred and nineteen students (35% of the eligible sample) responded at all five time points. The mean ages at T1, T2, T3, T4, and T5 were 28 (s.d. = 2.8), 29 (2.8), 31 (2.8), 37 (2.7) and 42 (2.7) years, respectively, with no significant sex differences. The mean observation times for T2, T3, T4 and T5 were 1.2 (s.d. = 0.2), 3.5 (0.3), 9.5 (0.6) and 14.6 (0.6) years after graduation (T1), respectively. Attrition analysis showed no significant associations between severe depressive symptoms, sex or age and the number of times of survey participation. Our sample of responders was not biased with respect to previous severe depressive symptoms, sex or age. See Fig. 1 for the study design.

<Please insert Figure 1 about here>

2.2. Measures

2.2.1. Dependent variables

Severe depressive symptoms were measured at all five observation points using the seven items that constitute the severe depression subscale of the General Health Questionnaire-28 (Goldberg and Williams, 1988). This questionnaire has been validated for screening clinically significant mental health problems in general population samples (Goldberg and Williams, 1988; Goldberg et al., 1998). Four of the seven items relate to suicidal ideation. The respondent is asked whether the symptoms have been present during the past 2 weeks, and the response categories range from (1) 'less than usual' to (4) 'much more than usual'. The values (1) and (2) were scored 0, and (3) and (4) were scored 1. Respondents with a sum score equal to or greater than 1 were considered 'severely depressed', whereas those scoring 0 were considered 'not severely depressed' (Vaglum and Falkum, 1999). The internal validity of our severe depressive symptom cut-off value of 1 was tested by calculating the overlapping responses with other clinically relevant categorical variables, viz. mental health treatment required and help-seeking (Tyssen et al., 2000), suicidal planning (Tyssen et al., 2004a) and impact on work capacity. The latter variable was assessed by one item: 'How often has your mental health made it difficult to master your work in the past 4 weeks?' This item was categorized from (1) 'not at all' to (5) 'all of the time' but was dichotomized to test for its association with severe depressive symptoms (1 was scored as 0; 2–5 were scored as 1).

Over the various observation points, 46%–87% of the subjects with severe depressive symptoms reported a need for mental health treatment compared with 8%–20% of subjects without symptoms; 13%–39% of students with severe depressive symptoms reported suicidal planning compared with 0%–6% of students without symptoms; and 10%–13% of students with severe depressive symptoms reported using antidepressant medications compared with

0%–3% of students without symptoms. At T5, 82% of subjects with severe depressive symptoms reported that their mental health condition had a negative effect on their work compared with 19% of subjects without symptoms (all χ^2: $p < 0.001$).

2.2.2. Predictor variables

Personality traits were measured in half the sample in the graduating semester (T1) and in the remaining half of the sample in the year after graduation (T2), as described elsewhere (Tyssen et al., 2000). Personality traits were measured using the 36-item version of Torgersen's Basic Character Inventory (Torgersen, 1980; Torgersen and Alnæs, 1989), which has been described previously (Tyssen et al., 2000). The *vulnerability* dimension closely resembles the classic neuroticism scale and will be referred to here as 'neuroticism'. It includes items such as 'I'm very touchy about criticism' and 'It often seems that others do things much better than me'. The *intensity* scale is a measure of extraversion/introversion, and the *control* dimension describes the degree of obsessiveness or conscientiousness. *Reality weakness* includes perceptions and ideations on the borderline between reality and fantasy. It measures chronic illusions, paranoid traits and problems with identity-insecurity and relationships, traits that are associated with severe personality disorders. Examples of these items are: 'It is difficult for me to trust people because they so often turn against me or leave me in the lurch', 'Sometimes I feel like I live in a fog' and 'Every now and then, I get strange thoughts in my head that I can't get rid of'.

Perceived parental bonding was measured in half the sample at T1 and in the remaining half at T2, for the same reasons that applied as to the measurement of personality traits (Parker et al., 1979). Perceived parental bonding was measured with the Parental Bonding Instrument, which asks respondents to rate the attitudes and behaviours of their parents on two dimensions: care and overprotection. Our research group has demonstrated

that maternal care has the strongest effect on depressive symptoms among doctors (Grotmol et al., 2010). Therefore, only this dimension was included in the analysis.

Perceived medical school stress was measured at T1 with an instrument that has been described previously (Vitaliano et al., 1984; Tyssen et al., 2000). This instrument consists of stressor items such as: 'Medical training controls my life and leaves too little time for other activities' and 'Medical school is more of a threat than a challenge'.

Perceived diagnostic skills were measured at T1 with an instrument constructed by our research group and described in detail elsewhere (Tyssen et al., 2000). This variable, ascertained with 16 items, describes how certain/uncertain the respondent feels about the way to approach a case history and a physical examination based on some specified medical diagnosis.

Perceived recording skills were measured at T1 by six items based on previous patient interviews and medical records (Tyssen et al., 2000). The items measured the respondent's confidence in his/her work during the preceding patient interviews and when medical records were written up during the previous term.

Hazardous drinking was measured at T1 with the Alcohol Use Disorder Identification Test (AUDIT) (Saunders et al., 1993). One of the ten items (item 2: 'How many drinks containing alcohol do you have on a typical day when you are drinking?') was omitted from the questionnaire based on the rationale that this item may be misunderstood in Norwegian settings. 'A typical day when you are drinking' may be interpreted as a typical day with a party or any other day when you are drinking alcohol (Gulbrandsen and Aasland, 2002). The present version has previously been used to study hazardous drinking among Norwegian doctors, police and ambulance officers (Sterud et al., 2007).

2.3. Statistical analysis

A preparatory analysis revealed that the rates of severe depressive symptoms were relatively unstable. Of the respondents who reported severe depressive symptoms at any observational point, 63.5% (35/56) reported them only once, 23.2% (13/56) reported them twice, 8.9% (5/56) three times, 3.6% (2/56) four times and 1.8% (1/56) reported severe depressive symptoms at all five observational points. Therefore, to capture a reliable measure of the respondents' tendency to experience severe depressive symptoms throughout the extensive follow-up period, we chose to assess the effects of the T1 variables on severe depressive symptoms at T2–T5 using generalized estimating equations (GEE). GEE is an extension of the generalized linear model that accounts for correlated repeated categorical measures within subjects (Diggle et al., 1994). To calculate the variance terms for repeated measures, the GEE approach requires that a working correlation structure be specified. An unstructured working correlation was specified for these analyses. The GEE approach uses all available data for each subject and is therefore less affected by missing data than are other statistical strategies.

To calculate easily comprehensible and meaningful risk estimates (odds ratios), an important aim of the study, we dichotomized the continuous predictor variables (except age), despite the potential loss of information entailed by the application of this strategy. The predictor variables were divided into groups of 'high' (at median or above) and 'low' (below median) scores. Because we expected severe depressive symptoms at T1 to be a strong predictor of future severe depressive symptoms, two predictor models were applied. The first adjusted model included all significant univariate predictors, except severe depressive symptoms at T1, whereas severe depressive symptoms at T1 was included in the second model.

The interactions between the different variables and time were tested to identify possible changes in the effects of the predictors over time. To identify any potential modifying effects of sex, interactions with sex were also tested. Finally, to test the possible modifying effects within pairs of variables, we performed two-way interaction analyses between all predictor variables. Differences in the prevalence of severe depressive symptoms between the five observational points were assessed by GEE.

Statistical significance was set at the 0.05 level.

2.4. Ethics

The study was approved by the Ethical Committee for Medical Research and the National Data Inspectorate of Norway.

Table 1 describes the independent variables.

<Please insert Table 1 about here>

3. Results

Severe depressive symptoms at T1 were reported by 13.7% of all subjects. There was a significant reduction ($p = 0.001$) in severe depressive symptoms of 7.7% from T1 to T2 and 7.2% from T1 to T5, which represent significant reductions ($p = 0.001$) from T1. There were no significant sex-based differences in the prevalence of severe depressive symptoms at any of the five observation points.

<Please insert Table 2 about here>

3.1. Univariate predictors of severe depressive symptoms

Table 2 shows that the following variables were significant univariate predictors of severe depressive symptoms: young age (OR 1.10 [1.02–1.19], $p = 0.01$), high neuroticism (OR 6.88 [3.42–13.81], $p < 0.001$), high reality weakness (OR 4.56 [2.56–8.14], $p < 0.001$), low maternal care (OR 2.64 [1.60–4.36], $p < 0.001$), high perceived medical school stress (OR 2.4 [1.4–4.0], $p = 0.001$), low perceived diagnostic skills (OR 2.15 [1.32–3.51], $p = 0.002$), low recording skills (OR 1.92 [1.18–3.12], $p = 0.01$) and severe depressive symptoms at T1 (OR 6.69 [4.09–10.94], $p < 0.001$).

3.2. Adjusted predictors of severe depressive symptoms

The following variables were significant adjusted predictors of severe depressive symptoms in the first multiple regression model: young age (OR 1.14 [1.04–1.24], $p = 0.01$), high neuroticism (OR 4.56 [1.88–11.04], $p = 0.001$), high reality weakness (OR 2.55 [1.27–5.10], $p = 0.008$) and low maternal care (OR 1.97 [1.17–3.32], $p = 0.01$). In the second model, in which we controlled for the severe depressive symptom status at T1, the following variables remained significant: young age (OR 1.12 [1.04–1.21], $p = 0.003$), high neuroticism (OR 3.40 [1.53–7.57], $p = 0.003$), high reality weakness (OR 2.28 [1.24–4.21], $p = 0.008$) and severe depressive symptoms at T1 (OR 3.62 [2.13–6.14], $p < 0.001$).

No interactions were identified with either time or sex, so that the other predictor effects were not dependent on sex or the observational point during the follow-up period.

4. Discussion

4.1. Main findings

The prevalence of severe depressive symptoms decreased from the graduating semester in medical school to the first postgraduate year and then stayed at about the same level until

the 15th postgraduate year. There were no differences in the prevalence of severe depressive symptoms according to sex. However, subjects with severe depressive symptoms or high scores for neuroticism or reality weakness in medical school ran a two- to fourfold increased risk of future severe depressive symptoms, and the risk conferred by reality weakness increased significantly when combined with hazardous drinking. Another novel finding was that younger medical students were at greater risk of future severe depressive symptoms. This is the first nationwide, long-term prospective study of doctors to estimate the early risk factors for severe depressive symptoms beyond the very first postgraduate years.

We observed a fourfold increased risk of future severe depressive symptoms among those who experienced these symptoms in medical school, which underscores the importance of early intervention among medical students to prevent future depressive episodes. It is also in keeping with a previous study that demonstrated that undergraduate depression is a long-term predictor of depression in male doctors (Firth-Cozens, 1998). However, to the best of our knowledge, the present study is the first to show this long-term effect among female doctors. Depression is not only devastating for the affected doctor, but can also interfere negatively with his/her patient care (Fahrenkopf et al., 2008). However, partly because of the high stigmatization of depressive illness, medical students and doctors are less likely than the general population to seek and receive appropriate treatment (Bramness et al., 1991; Töyry et al., 2000). Therefore, it must be ensured that they receive early and proper treatment. Because rates of suicidal ideation are elevated in this group (Dyrbye et al., 2008), some doctors may also benefit from antidepressant medication. From our clinical experience, we have reason to believe that bipolar disorder type II is relatively prevalent in this occupational group (Gabbard and Myers, 2008). The relatively low stability of the depression measure observed in our sample indicates fluctuations in mood levels over the years, as seen in bipolar disorder, but this conjecture requires further study.

Young age predicted severe depressive symptoms. This is an important finding, and should direct attention to the youngest students, who have less experience in both work and relationships and who are exposed to stressful medical education and a subsequent stressful career. Until recently, one could apply for admission to Norwegian medical schools after earning extra points through work experience or by improving high school grades. Today, students commonly enter medical school directly after high school (aged 19 years), so they are younger than the present study sample and may be even more vulnerable to severe depressive symptoms. These considerations may also apply to other countries that enrol quite young students. Theories of psychological development suggest that identity formation is still ongoing in the early twenties (Lerner, 2002), and our findings suggest that some young adults encounter the challenges of medical school somewhat prematurely. For instance, demanding and time-consuming medical school studies may negatively affect their personal life by impeding the development of social skills and social support (Kjeldstadli et al., 2006). These findings suggest that the admission age to medical school should be reconsidered. Would it be wise to delay entrance to the overwhelming medical curriculum for a couple of years?

We found that reality weakness was an independent risk factor for subsequent severe depressive symptoms. To some extent, this is consistent with the finding that personality psychopathology increases the risk of suicidal ideation (Brezo et al., 2006). As reality weakness also predicts a lack of help-seeking, independent of the level of emotional distress (Tyssen et al., 2004b), the medical faculty and colleagues of doctors in trouble should be aware that this trait might be a marker of severe future problems and even serious suicidal ideation (Tyssen et al., 2004a). Contrary to studies of other populations, which have claimed that alcohol disorders increase the risk of depression (Fergusson et al., 2009), we found no significant predictive effect of hazardous drinking in its own right. The finding that high neuroticism predicts depression concurs with previous (Tyssen and Vaglum, 2002) and recent

(Sen et al., 2010) studies of interns. The effect of neuroticism was only slightly affected by severe depressive symptoms in the graduating semester, illustrating that neuroticism and severe depressive symptoms are distinct concepts. Therefore, self-criticism and neuroticism should be targeted in psychotherapy (Jorm, 2000). For preventive purposes, the risks associated with these traits and the aggravating role of hazardous drinking should also be taught to students as part of their curriculum.

Studies of general population samples have reported that depressive symptoms tend to decrease from early adulthood to mid-life (Jorm, 2000), and the decline in the prevalence of severe depressive symptoms observed in the present study is consistent with this. There are discrepancies between previous findings among medical students and young doctors with regard to sex-based differences: some studies have reported no differences, whereas the majority have found more depression among women (Tyssen and Vaglum, 2002; Grotmol et al., 2010). The lack of sex-based differences in the present study may be attributable to: (1) the selection of female students with good mental health; (2) the selection of male students with poor mental health; or (3) both.

4.2. Strengths and limitations

The strengths of this study include the nationwide sample, the longitudinal design and the application of a widely used measure to capture *severe depressive symptoms*, which has also been confirmed to be of clinical importance. Because the rates of severe depressive symptoms were unstable over time, we attempted to establish a reliable measure of severe depressive symptoms over time by assessing the population on five occasions. The possible effects of missing data were also considerably reduced by using a GEE approach.

We are aware that contextual factors, such as work stress and negative life events, can influence depressive symptoms, and further longitudinal research should clarify such co-

occurrent factors. There are obvious limitations associated with the use of self-report measures, but these were counteracted, to some extent, by controlling for neuroticism in our predictor model, because this trait resembles negative affectivity (Depue and Monroe, 1986). However, medical students may also be reluctant to express their vulnerability (Chew-Graham et al., 2003), and the effects documented in our results could be underestimated. Although the influence of missing data was reduced in this study, the non-response bias may still have affected the results. However, an attrition analysis found no greater prevalence of severe depressive symptoms in the graduating semester among late responders, which tends to validate our findings.

Severe depressive symptoms in medical school are a risk factor for future severe depressive symptoms, as are young age and particular personality traits. Neuroticism and reality weakness can be treated with psychotherapy. Finally, medical students who suffer from severe depressive symptoms should be identified and offered the appropriate treatment, perhaps with extra attention paid to the youngest students.

References

Bramness, J.G., Fixdal, T.C., Vaglum, P., 1991. Effect of medical school stress on the mental health of medical students in early and late clinical curriculum. Acta Psychiatr. Scand. 84, 340–345.

Brezo, J., Paris, J., Tremblay, R., Vitaro, F., Zoccolillo, M., Hébert, M., Turecki, G., 2006. Personality traits as correlates of suicide attempts and suicidal ideation in young adults. Psychol. Med. 36, 191–202.

Center, C., Davis, M., Detre, T., Ford, D.E., Hansbrough, W., Hendin, H., Laszlo, J., Litts, D.A., Mann, J., Mansky, P.A., Michels, R., Miles, S.H., Proujansky, R., Reynolds III, C.F., Silverman, M.M., 2003. Confronting depression and suicide in physicians: a consensus statement. JAMA 289, 3161–3166.

Chandavarkar, U., Azzam, A., Mathews, C.A., 2007. Anxiety symptoms and perceived performance in medical students. Depress. Anxiety 24, 103–111.

Chew-Graham, C.A., Rogers, A., Yassin, N., 2003. 'I wouldn't want it on my CV or their records': medical students' experiences of help-seeking for mental health problems. Med. Educ. 37, 873–880.

Depue, R.A., Monroe, S.M., 1986. Conceptualization and measurement of human disorder in life stress research: the problem of chronic disturbance. Psychol. Bull. 99, 36–51.

Diggle, P.J., Liang, K.Y., Zeger, Z.L., 1994. Analysis of Longitudinal Data. Clarendon Press, Oxford, England.

Dyrbye, L.N., Thomas, M.R., Shanafelt, T.D., 2006. Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students. Acad. Med. 81, 354–373.

Dyrbye, L.N., Thomas, M.R., Massie, F.S., Power, D.V, Eacker, A., Harper, W., Durning, S., Moutier, C., Szydlo, D.W., Novotny, P.J., Sloan, J.A., Shanafelt, T.D., 2008. Burnout and suicidal ideation among U.S. medical students. Ann. Intern. Med. 149, 334–341.

Fahrenkopf, A.M., Sectish, T.C., Barger, L.K., Sharek, P.J., Lewin, D., Chiang, V.W., Edwards, S., Wiedermann, B.L., Landrigan, C.P., 2008. Rates of medication errors among depressed and burnt out residents: prospective cohort study. Br. Med. J. 336, 488–491.

Fergusson, D.M., Boden, J.M., Horwood, L.J., 2009. Tests of causal links between alcohol abuse or dependence and major depression. Arch. Gen. Psychiatry 66, 260–266.

Firth-Cozens, J., 1998. Individual and organizational predictors of depression in general practitioners. Br. J. Gen. Prac. 48, 1647–1651.

Gabbard, G.O., Myers, M.F., 2008. The Physician as a Patient: A Clinical Handbook for Mental Health Professionals. American Psychiatric Publishing Inc., Washington, DC.

Goldberg, D., Williams, P., 1988. A User's Guide to the General Health Questionnaire. NFER-Nelson, Windsor.

Goldberg, D., Privett, M., Ustun, B., Simon, G., Linden, M., 1998. The effects of detection and treatment on the outcome of major depression in primary care: a naturalistic study in 15 cities. Br. J. Gen. Prac. 48, 1840–1844.

Grotmol, K.S., Ekeberg, Ø., Finset, A., Gude, T., Moum, T., Vaglum, P., Tyssen, R., 2010. Parental bonding and self-esteem as predictors of severe depressive symptoms: a 10-year follow-up study of Norwegian physicians. J. Nerv. Ment. Dis. 198, 22–27.

Gulbrandsen, P., Aasland, O.G., 2002. Changes in drinking habits among Norwegian physicians 1985–2000. Tidsskr. Nor. Laegeforen. 122, 2791–2794.

Hem, E., Haldorsen, T., Aasland, O.G., Tyssen, R., Vaglum, P., Ekeberg, O., 2005. Suicide among physicians. Am. J. Psychiatry 162, 2199–2200.

Jorm, A.F., 2000. Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult life span. Psychol. Med. 30, 11–22.

Kjeldstadli, K., Tyssen, R., Finset, A., Hem, E., Gude, T., Gronvold, N.T., Ekeberg, O., Vaglum, P., 2006. Life satisfaction and resilience in medical school—a six-year longitudinal, nationwide and comparative study. BMC Med. Educ. 6, 48.

Lerner, R.M., 2002. The differential approach. In: Webber, B. (Ed.), Concepts and Theories of Human Development. Lawrence Erlbaum Associates Inc., Mahwah, pp. 409–437.

Parker, G., Tupling, H., Brown, L.B, 1979. A parental bonding instrument. Br. J. Med. Psychol. 52, 1–10.

Saunders, J.B., Aasland, O.G., Babor, T.F., de la Fuente, J.R., Grant, M., 1993. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption—II. Addiction 88, 791–804.

Schernhammer, E.S., Colditz, G.A., 2004. Suicide rates among physicians: a quantitative and gender assessment (meta-analysis). Am. J. Psychiatry 161, 2295–2302.

Sen, S., Kranzler, H.R., Krystal, J.H., Speller, H., Chan, G., Gelernter, J., Guille, C., 2010. A prospective cohort study investigating factors associated with depression during medical internship. Arch. Gen. Psychiatry 67, 557–565.

Sterud, T., Hem, E., Ekeberg, O., Lau, B., 2007. Occupational stress and alcohol use: a study of two nationwide samples of operational police and ambulance, personnel in Norway. J. Stud. Alcohol Drugs 68, 896–904.

Torgersen, S., 1980. Hereditary–environmental differentiation of general neurotic, obsessive, and impulsive hysterical personality traits. Acta Genet. Med. Gemellol. (Roma) 29, 193–207.

Torgersen, S., Alnæs, R., 1989. Localizing DSM-III personality disorders in a three-dimensional structural space. J. Pers. Disord. 3, 274–281.

Töyry, S., Räsänen, K., Kujala, S., Aärimaa, M., Juntunen, J., Kalimo, R., Luhtala, R., Mäkelä, P., Myllymäki, K., Seuri M., Husman, K., 2000. Self-reported health, illness, and self-care among Finnish physicians: a national survey. Arch. Fam. Med. 9, 1079–1085.

Tyssen, R., Vaglum, P., Gronvold, N.T., Ekeberg, O., 2000. The impact of job stress and working conditions on mental health problems among junior house officers. A nationwide Norwegian prospective cohort study. Med. Educ. 34, 374–384.

Tyssen, R., Vaglum, P., 2002. Mental health problems among young doctors: an updated review of prospective studies. Harv. Rev. Psychiatry 10, 154–165.

Tyssen, R., Hem, E., Vaglum, P., Gronvold, N.T., Ekeberg, O., 2004a. The process of suicidal planning among medical doctors: predictors in a longitudinal Norwegian sample. J. Affect. Disord. 80, 191–198.

Tyssen, R., Rovik, J.O., Vaglum, P., Gronvold, N.T., Ekeberg, O., 2004b. Help-seeking for mental health problems among young physicians: is it the most ill that seeks help? A longitudinal and nationwide study. Soc. Psychiatry Psychiatr. Epidemiol. 39, 989–993.

Vaglum, P., Falkum, E., 1999. Self-criticism, dependency and depressive symptoms in a nationwide sample of Norwegian physicians. J. Affect. Disord. 52, 153–159.

Vitaliano, P.P., Russo, J., Carr, J.E., Heerwagen, J.H., 1984. Medical school pressures and their relationship to anxiety. J. Nerv. Ment. Dis. 172, 730–736.

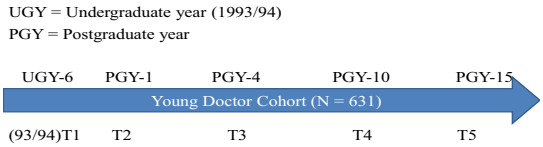


Fig. 1. Study design.

Table 1

Description of the independent variables with cut-off thresholds.

	n	Range	Mean (s.d.)	Median (used for cut-off)
Age (years)	505	24–49	28 (2.83)	
Sex	522		57% females	
Neuroticism	459	0–9	3.49 (2.31)	3
Reality weakness	459	0–9	1.09 (1.42)	1
Intensity	459	0–9	5.56 (2.46)	6
Control	459	0–9	3.05 (2.03)	3
Perceived medical school stress	514	3–40	20.35 (6.85)	20
Perceived diagnostic skills	516	42–110	85.57 (10.24)	86
Perceived recording skills	517	15–42	28.64 (4.75)	29
Maternal care	451	0–36	28.48 (6.52)	30
Hazardous drinking	346	0–40	3.02 (2.60)	3

Table 2

Risk of future severe depressive symptoms (T2–T5).

	Univariate		**Adjusted**			
			Model 1		Model 2	
	OR (95% CI)	*p* value	OR (95% CI)	*p* value	OR (95% CI)	*p* value
Young age	1.10 (1.02–1.19)	0.01	1.14 (1.04–1.24)	0.01	1.12 (1.04–1.21)	0.003
Sex	1.05 (0.65–1.72)	0.84	1.30 (0.73–2.31)	0.37	1.19 (0.72–1.98)	0.5
Neuroticism	6.88 (3.42–13.81)	0.001	4.56 (1.88–11.04)	0.001	3.40 (1.53–7.57)	0.003
Reality weakness	4.56 (2.56–8.14)	0.001	2.55 (1.27–5.10)	0.008	2.28 (1.24–4.21)	0.008
Intensity	0.92 (0.57–1.50)	0.74	[a]	[a]	[a]	[a]
Control	1.16 (0.71–1.89)	0.56	[a]	[a]	[a]	[a]
Perceived medical school stress	2.4 (1.4–4.0)	0.001	0.96 (0.53–1.74)	0.89	[c]	[c]
Low perceived diagnostic skills	2.15 (1.32–3.51)	0.002	1.79 (0.97–3.32)	0.06	[c]	[c]
Low recording skills	1.92 (1.18–3.12)	0.01	0.98 (0.54–1.79)	0.95	[c]	[c]
Low maternal care	2.64 (1.60–4.36)	0.001	1.97 (1.17–3.32)	0.01	1.49 (0.90–2.49)	0.13
Hazardous drinking	1.73 (0.96–3.11)	0.07	[a]	[a]	[a]	[a]

Severe depressive symptoms at T1	6.69 (4.09–10.94)	0.001	[b]	[b]	3.62 (2.13–6.14)	0.001

OR = Odds ratio

CI = Confidence interval

[a] Variable not significant in the univariate analysis, therefore not included in the adjusted analysis.

[b] Variable not included in the first model.

[c] Variable not significant in the first model, so was not included in the second model.